UNIVERSITY OF ALBERTA

Alternative Assessment in Senior High School Classrooms

BY

ROBERT J. RITTER ©

A thesis submitted to the Faculty of Graduate Studies and research in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

DEPARTMENT OF SECONDARY EDUCATION

FALL 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-18100-6

Canadä

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF THE AUTHOR: ROBERT J. RITTER

TITLE OF THESIS: ALTERNATIVE ASSESSMENT IN SENIOR HIGH SCHOOL SCIENCE CLASSROOMS

DEGREE: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE WAS GRANTED: 1996

Permission is hereby granted to the University of Alberta to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except hereinbefore provided the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Robert Ritter
15271 43 avenue
Edmonton Alberta
T6H 5R3

May 24/96

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for the acceptance, a thesis entitled ALTERNATIVE ASSESSMENT IN SENIOR HIGH SCHOOL SCIENCE CLASSROOMS here submitted by ROBERT J. RITTER in partial fulfillment of the requirement of the requirements for the degree of DOCTOR OF PHILOSOPHY.

Dr. Wallie Samiroden

Jim Parson

Dr. James Parson

Dr. Wytze Brouwer

Dr. Ken Jacknicke

G. D. Taylor

Dr. Gerald Taylor

Dr. Gaalen Erickson

May 6, 1996

Abstract

This research study provides a profile of three independent, action research groups involved in seeking alternative assessment strategies for senior high school science classrooms. The direction for research evolved spontaneously from inservice work, designed to help teachers at three different high schools cope with provincial curriculum changes mandated for September 1992. Teachers identified current assessment approaches as insufficient in supporting the intent of new curriculum initiative.

The action research project was an attempt to begin asking questions about why current assessment practices exist. As a researcher I began by exposing the hidden and not so hidden assumptions. The movement toward portfolios, journals, and performance-based assessment allowed students to become involved in decision-making. Breaking down a structure that links the notion of "expert" to power and provides authority, requires that students and teachers begin viewing assessment in a different way. Assessment is not something that teachers do to students, but a way of coming to understand learning.

A profile of nine teachers involved in the action research study along with the ecology of the three different high schools is provided. Northern High School provides an account of a successful research project in a new location, staffed by a predominantly young and energetic staff. Eastern High School provides a description of a slower moving, but successful research project, in an older school, staffed by teachers with stronger links to tradition. Southern High School profiles an unsuccessful project, where a dynamic group of dedicated teachers attempted to bring about changes in assessment practices while struggling to adjust to Site-Based-Management and meet other increasing demands placed upon the staff.

The dichotomy of curriculum intent and assessment strategies continues to exist. If anything the perceived gap is better understood and possibly metaphorically wider. What was accomplished, by action research groups, was providing a forum for initiating discussions about

changing assessment strategies with colleagues, and a way of including students in determining what they have learned.

Table of Contents

Chapter 1: Outline of the Research Question

Chapter 1: Outline of the Research Question 1
- Overview 1
- Identification of a problem 1
- The research questions 2
- Description of the study 3
  - Chronology of organizational tasks: 4
  - Launching action research 4
  - Extending the research to include students 5
  - Developing templates for alternative assessment 6
- Methodology for data gathering 7
- Context for the study 8
  - Northern High School 8
  - Southern High School: 9
  - Eastern High School 10
- Context for the interviews 10
  - Teacher Interviewees 11
  - In-depth profile of teacher interviewees 12
- Significance of the study 15
  - Delimitation 16
  - Limitations 17
  - Assumptions 18
- Definition of terms 18

Chapter 2: The Rationalist Model for Assessment 21
- Overview 21
- Influence of the Cartesian rationalism 21
  - Notion of the superiority of objective knowledge 24
  - Method and power 26
  - To arrogate method 28
- Rationalism as a method for education 31
  - Limitations of the standardized testing model 32
  - The distortion of repeatability and certainty 33
  - The humanist alternative 35
  - Notion of connoisseurship 38
- Criteria for authenticity 38

Summary table of traditional and alternative assessment strategies 40
Traditional view and expectations 40
Emerging view and expectations 40
Finding a direction for change 40
Obstacles to change 41
The movement toward alternative assessment: England 43
The movement toward alternative assessment: Canada 46
Summary 49
Chapter 3: Adopting a research approach 50
Overview 50
Research Orientations 50
The selection of the action research orientation 54
Describing the action research orientation 54
What action research is not: 56
Collaborative professional development through action research 56
Finding an action research theme 58
Collaborative change 58
Change from within the school 59
Origins of the action research groups 60
Summary 62
Chapter 4: Evolution within the Action Research Groups 64
Overview 64
Reflection upon the action research conducted at Northern school 64
A. Initiation phases 66
B. Development phases 68
C. Reconnaissance phases 70
Summary observations and interpretations 70
Reflection upon the action research conducted at Eastern School 71
A. Initiation phases 73
B. Development phases 75
C. Reconnaissance phases 77
Summary observations and interpretations 77
Reflection upon the action research conducted at Southern School 78
A. Initiation phases 80
B. Developmental phases 81
Summary of observations and interpretations 82

Chapter 5: Documenting changes in assessment strategies 84
    Overview 84
    Identification of problems with traditional laboratory reports 84
        Portfolios and revising answers 85
        Policing copied work and the nature of science 87
    Students abdicate ownership 89
        Portfolios and student ownership 90
        Hidden consequences of accountability 94
    Teachers' difficulties in abandoning old paradigms 96
        Teacher concerns about changing assessment strategies 97
        Student perceptions about teacher concerns 98
    Benefits to students 101
    Summary 102
Chapter 6: Conclusions and Implications 104
Overview 104
    Reflecting upon the research questions 104
        Discordance of intents 104
        Support for changing assessment strategies 107
        Determining what is valued 108
        Using teacher practical knowledge 109
        Factors that limit change 110
        Problems created by assessing a broader curriculum 112
    Unanswered Questions 115
    Conclusion 117
    Bibliography 120
Thesis Appendices 126
    Appendix #1: Faculty of Education Ethics Review 126
    Appendix #2: Cooperative Activities Program 128
    Appendix #3 Parent approval letter 130
    Appendix #4: Sample student interview questions 130
    Appendix #5: Students' Questionnaire Portfolio Assessment 132
    Appendix #6: Assessment Templates 133
        6A: Assessment template for Peer Assessment (group) 133
        6B: Assessment Template for Group Work 134
        6C: Assessment Template for Debates 135
        6D: Assessment Template for Presentation of Research Topics 136

6E: Assessment Template for a Teacher-Directed Laboratory Investigation 137
6F: Assessment Template for a Student-Designed Laboratory Investigation 138
6G: Assessment Template for Formative Assessment of Student Problem-Solving 139
6H: Assessment Template for Scientific Inquiry 140
6I: Assessment Template for Biological Drawings 141
6J: Portfolio Assessment: Science 10 Unit 2, 1992 142
6K: Scoring Rubric for Portfolios: Science 10 (unit 2), 1992 144
6L: Sample Portfolio Assignment: Biology 30, Term #1, 1992 145
6M: Scoring Rubric for Portfolios: Biology 30, 1992 146
Appendix #7: Lessons developed 147
7A: Investigation: Using a Microscope for Forensic Investigations 147
7B: Investigation: Writing a Mystery 149
Appendix #8: Journal Entries 154
Appendix 9: Performance-Based Assessment Strategies 157
9A: Science 10: Microscope Work 157
9A: Assessment Template for the Microscope 157
9B: Science 10, Osmosis and the Concentration of Solute 159
9B: Science 10, Performance-Based-Assessment for Challenge 160
9C: Science 14, Light Unit 161
9C: Science 14, Performance-Based-Assessment 162
9D: Biology 30, Dynamic Model of the Cell 163
9D: Biology 30, Performance-Based-Assessment for The Dynamic Model of Cell Division 163
9E: Biology 20, Designing an Experiment to Determine How Environmental Factors Affect Seed Germination 165
9E: Biology 20, Assessment of Student Performance 167
9F: Chemistry 30: Antioxidants 169
9F: Chemistry 20, Assessment Template Antioxidants 169
9G: Chemistry 20, Determining Vitamin C Levels in Juice Drinks 170
9G: Chemistry 20, Assessment Template Determining Vitamin C Levels in Juices and Drinks 171
10: Table of inventions 172

Table of Figures

Figure #1: Research questions 50
Figure #2 Kurt Lewin's view of action research: 54
Figure #3: Elliot's action research spiral 55

Chapter 1: Outline of the Research Question

Overview

In this chapter I will recount the research questions that arose, as three action research groups worked independently to provide assessment strategies that support curriculum change. In that many of the research questions were those posed by teachers, I have purposely separated research questions of interest to practicing teachers and those questions which appealed to me as a researcher. The chapter provides an outline of the research methodology and profiles of research sites and teachers involved in the study.

Within the chapter, I will described the significance of the study. Delimitations are described and limitations are acknowledged in describing the intents of the action research groups. In helping to provide a context for the study, I have provided a list of commonly used terms for the reader. In part, these terms will serve as a reference to my research orientation and provide a framework for developing a philosophical framework for understanding the questions posed.

Identification of a problem

The gaps between theory and practice resonate in the tensions that exist between curriculum as planned, curriculum as implemented, and curriculum as learned. Nowhere is the gap more evident than in the apparent dichotomy of curriculum intents and the countenance of student assessment. The 1992 implementation of new science curricula for Alberta with its broadened goals, subsumed in what has become referred to as science-technology-society or STS science, has magnified the discordance between the goals for instruction and student assessment.

An STS approach to curriculum is designed to provide a social context for learning science by focusing on the social impact brought about by the interaction of science and technology. No longer is science to be presented as a compilation of decontextualized facts and pre-made theories. The lived experiences of students must be recognized as programs attempt to become personally meaningful and socially relevant. Alberta Education's "Vision for High School Science Programs" states:

> The senior high school programs will help all students attain the scientific awareness needed to function as effective members of society. (Alberta Education, 1991, p. 1)

The vision statement goes on to indicate:

> Students will be expected to show an appreciation for the roles of science and technology in understanding nature and maintaining a lifelong interest in science.

> They will possess positive attitudes toward and enthusiasm for science (1991, p. 1)

Alberta Education indicates that the achievement of these goals is dependent upon creating a context for learning. They state:

> The learning opportunities will be made meaningful by providing concrete experiences that students relate to their world. (1991, p. 1).

The call for a context to relate school science with life experiences is linked with a call for greater responsibility for student learning. Phrases such as "Students will be expected to show an appreciation for" and "They will possess" underscores a political climate that demands accountability.

Although few science teachers would oppose the intent of a science program that looks beyond the confines of acquiring scientific knowledge, confusion abounds as how to provide a program that ensures that "all students" attain the scientific awareness necessary to guarantee that they become "effective members of society," or to provided the surety that students develop "an appreciation for the roles of science and technology." Inherent within the statements are the a priori assumptions that a self-evident definition for "scientific awareness" and "effective members of society" exists or that a methodology for developing positive attitudes waits to be discovered. If the statement is interpreted with purpose rather than the cynicism that envelops banal political rhetoric, teaching science and assessing learning must be viewed in a new light.

The introduction of environmental case studies, role-playing scenarios, interactive debates, the sharing of personal narrative about science, and student-designed laboratory procedures that stress technological-thinking-skills, has received greater attention among those who call for a change in the manner in which science is presented (Aikenhead, 1980; Risi, 1982; Baird, 1986; Tobin, Espinet, Byrd, & Adams, 1988; ; Geddis, 1991; Wolf, Bixby, Glenn, & Gardner,1991; Martin, 1991; Villani, 1992). In addition, a perceived division exists between the manner in which students are currently being assessed and the manner in which curriculum is intended to be presented (Eisner, 1985; Lewis, 1992; Wiggins, 1992b). An evaluation model grounded within the technocratic principles of standardized assessment, provides no or little evidence of looking at positive attitudes (Aoki, 1984; Eisner, 1985; Wiggins, 1989; Wiggins, 1992a; Wiggins, 1992b; Wolf, Bixby, Glenn, & Gardner, 1991) or addressing the social context in which science develops (Collins, 1992; Geddis, 1991; Tobin, Butler-Kahle, Fraser, 1991).

The research questions

The following questions were proposed prior to the study:

1. Are the evaluation strategies being used by high school science teachers to assess students' knowledge of science compatible with the intents of the new curriculum?
2. Have assessment approaches changed?
3. Can an action-research group develop assessment strategies that complement STS teaching strategies? What components can be evaluated? What message will we send to students if we don't evaluate a certain component of STS? Will the students value components of the course not evaluated?
4. What problems will be created as we attempt to move toward assessing technological thinking and social issue components of a science program?

As a researcher, a number of questions, many unanswered, evolved during the study:

1. Can teachers work in a collaborative action research setting to develop alternative assessment strategies? What impact will a collaborative action research project, involving three schools, have upon the teachers, students, administrators, and communities associated with each of the schools?
2. What are the teachers' reasons, within the framework of teacher practical knowledge (Connelly, 1985; Duffee & Aikenhead, 1992) or personal knowledge (Polanyi, 1962), for including or excluding assessment practices that could evaluate student knowledge of the three STS components: nature-of-science, science-and-technology, and science-and-society?
3. Can an action research group bridge the gap between curriculum intent and evaluation practices (Casonova, 1989; Dorr-Breme, 1983; Eisner, 1985) or should we attempt to bridge the gap? Does the institutional setting (because of pressures from the School District, Alberta Education, or the community) advocate goals inconsistent with developing aspects of STS curriculum and evaluation?
4. Do students agree with teacher perceptions of student needs concerning assessment practices (Duffee & Aikenhead, 1992; Herman, 1992b)?

Description of the study

An action research project was preceded by a pilot study. The pilot study arose serendipitously from science 10 workshops with St. Albert schools. A school from the pilot study, identified as Northern High School, was included in the main study. The main research study began in September 1993 and also included two Edmonton Catholic Schools, identified as Eastern and Southern High Schools.

Chronology of organizational tasks:

- Ethics review completed and permission to be obtained for pilot study Jan. 1993. See appendix#1.
- Cooperative Activities form, completed for Field Experiences outlining the nature of the study, Jan. 1993. See appendix #2.
- A pilot action research group consisting of teachers from three high schools in St. Albert was established in February 1993.
- Permission obtained from central office personnel for Edmonton Catholic Schools and St. Albert Protestant Schools, February 1993.

Launching action research

- A series of organizational meetings were conducted with interested teachers from the pilot schools. A common focus was established; developing assessment strategies that supported the teaching of science within an STS context. February 1993 for Northern High School, and June 1993 for Southern and Eastern High Schools.
- A series of weekly meetings were established for the first month of the project, followed by monthly meetings thereafter. I attended each of the weekly meetings at Northern and Eastern High School, but was only able to attend 2 of the four initial meetings held at Southern High School. I attended monthly meetings held at Northern and Eastern High Schools until the completion of the project. Meetings at Southern were often rescheduled. I was only able to attend two monthly meetings. Most action research meetings lasted approximately 45 minutes.
- No chairperson was appointed for the action research meetings, nor was an agenda established prior to the meeting. Teachers spoke freely about classroom experiences with alternative assessment or set times to work on cooperative planning. In many meetings, I was asked to share what was happening in other groups.
- Working groups, at each of the high schools, developed alternative assessment strategies. I prepared contact summaries (Miles & Huberman, 1984) following each of the meetings identifying main themes, target research questions and the impressions of members of the action research group. June 1993 and November 1993 for Northern High School, and December 1993 and May 1994 for Southern and Eastern High Schools.
- A description of alternative assessment strategies was taken from the Teacher's Edition of *Nelson Biology* (Ritter & Samiroden, 1993), and *A Model for the Assessment and Evaluation of Scientific Problem-Solving Skills* (Alberta Education, 1992) was provided to initiate discussion. February 1993 for Northern High School, and June 1993 for Southern and Eastern High Schools. A two

hour orientation seminar was conducted for all members of the science staff at each school, after teachers had an opportunity to read the articles. During the orientation seminar, teachers had the option to become involved in the project.

- Although all members of the science departments at each of the school were invited to become involved with the action research study, nine teachers were selected for in-depth interviews. Two interviews were audio taped and notes were prepared from the interviews. Researcher impression and interpretations were checked with interviewees. February 1993 for Northern High School, and June 1993 for Southern and Eastern High Schools.
- Teacher interviewees were asked to keep a log book of their impressions about how new assessment strategies are perceived by students, their colleagues, and themselves. Two of the nine teachers provided complete entries, the other seven moved away from the log book entries as assessment practices progressed. Teacher impressions, recorded in point form, were designed to act as an outline to construct interview questions. Teacher validation was sought for all interpretations from log books. All teacher journals were returned to the teachers in June 1995, as agreed to prior to beginning the research project.

Extending the research to include students

- Selected student interviews were conducted in groups of two or three. All sixteen interviewees were recorded on audio tape. Eight students were interviewed in the spring of 1993 and another eight students were interviewed during the fall and winter of 1994. Two interviews were conducted with each student group. Because interviews were audio taped, student identity was protected and parental permission for each interview was obtained. (See appendix #4 for sample student interview questions. See appendix #3 for parent's approval letter). The questions presented are intended to indicate the direction of the interview, rather than to be an inclusive list of sequenced questions. Student responses and discussion helped structure the interviews. Teacher perceptions of student's needs were checked during the student interviews.
- Students from six different classes (n= 155) at Eastern High School were surveyed, during the 1993/94 school year, for their impressions about the alternative assessment strategies attempted. The survey was conducted in June 1994 for all members of the class. Computerized scoring cards were used and student names were not placed on the cards. Although no attempt was made to identify individual students, each class was identified by a code number on the scoring card. Students surveyed came from Science 10, Biology 20 and Biology 30 classes and were taught by three different teachers. An

item analysis was conducted on survey results. See student questionnaire questions and results in appendix #5.

Developing templates for alternative assessment

- I worked with teachers at Eastern High School to prepare a series of sample alternative assessment strategies for Science 10 , Biology 20, and Biology 30. The list includes portfolios (see appendix #6: 6 J to 6 M), metacognitive journals taken from biology 20 and 30 (see appendix #8), performance-based laboratory assessment on the microscope for science 10 (see appendix 9A), and a murder mystery for science 10 that employs peer assessment and personal reflection of group performance. (See appendices 7A and 7B, for the Murder Mystery and appendices 6A and 6B for the assessment templates for peer assessment and group work). The assessment templates were made available to teachers at all sites regardless of whether they attended action research meetings between February 1993 and January 1994.
- The action research groups at Northern and Eastern high schools developed lessons that employed performance-based assessment for Science 10, Science 14, Biology 20, Chemistry 20, and Biology 30. (See appendices 9B to 9G). The lessons were developed between February 1993 and June 1994.
- The action research group at Eastern high school developed generic assessment templates for debates (see appendix 6C), the presentation of research topics (see appendix 6D), the assessment for teacher-directed laboratory investigations (see appendix 6E), the assessment for student-designed laboratory investigations (see appendix 6F), the formative assessment of problem solving (see appendix 6G), the assessment of scientific inquiry (see appendix 6H),and the assessment of biological drawing (see appendix 6I). The assessment templates were developed between February 1993 and June 1994.

Methodology for data gathering

1. Interviews:

Type	Location	Process
Teachers (9) Small group interview (groups of 2 or 3)	Northern High School (2 teachers) Southern High School (2 teachers) Eastern High School (5 teachers)	• Interview conducted after teachers used portfolio assessment for two months. • Questions taken from teacher's log book. • 20 minute interview • Validation of interpretations from action research meetings.
Teacher (9) Individual interviews	Northern High School (2) Southern High School (2) Eastern High School (5)	• Interview conducted 10 months after dealing with alternative assessment. • Reaction to other teachers' impressions and student interview comments. • Validation of interpretations from previous action research meetings.
Students (16)	Small group (2/3), location Eastern High School	• See appendix #4 for list of questions • Reaction to teacher's impressions.

2. Student survey (n=155) dealing with student perceptions of alternative assessment strategies attempted at Eastern High school. The survey involved students from 6 different classes, taught by three different teachers. Two classes were at the grade 10, grade 11, and grade 12 levels respectively.
3. Interim summaries (Miles & Huberman, 1984) were prepared in June 1993, for the pilot study, and during December 1993 and during May 1994 for the main study. The interim summaries synthesized my thinking and provided direction for reflection.
4. Contact summaries that follow action research meetings (Miles & Huberman, 1984). My contact summaries were not shared with other teachers in the action research group. The summaries provided an outline of my reactions to the meetings that were designated as sharing sessions both for assignments that are intended to develop alternative assessment strategies and the marking schemes that were developed to evaluate student

work. Teachers were informed that summaries would be kept and assured that anonymity would be preserved.

5. I prepared a researcher journal in which I recorded my interpretation of changes in my own classroom as alternative assessment strategies were implemented. The Journal was kept between February 93 until January 1994.
6. A student questionnaire was prepared for six classes of students at Eastern high school during the spring of 1994. The classes were at three different grade levels (grade 10, 11, and 12 all academic classes) and from three different teachers. (See appendix #5 for the questionnaire and appendix #6 for student results).
7. Student journal entries and portfolios from my classes and selected samples of those collected by other teachers at Eastern High School were also used for research purposes. Four of the teachers at Eastern High School frequently shared student portfolios and engaged in some cross-instructor portfolio marking.

Context for the study

One of the St. Albert Schools (identified as Northern High School) and two Edmonton Catholic Schools (identified as Eastern and Southern High Schools) were used in the study. The study at all schools involved participation by more teachers than were interviewed. At every high school, some teachers gradually moved away from the action research group for a variety of reasons. The teachers chosen for interviews at Northern and Southern high school became the informal leaders in their science department and of their action research group. The research at Eastern High school provided the greatest focus for the study. This occurred largely because I was teaching at the school and, therefore, extended research opportunities were readily available.

Northern High School

The Northern High School participated in the pilot study. I believed Northern High School to be the most active research group of the pilot sites. In part, the success originated with the department head, who was especially interested in alternative assessment. The Northern High School group is composed of experienced teachers in a newer school. The community is upper middle class but the school has a mix of academic and vocational courses. Northern High School provided a collegial atmosphere, new setting, and evolving school philosophy for the action research group. The school administration and central office administration became interested in action research. Although neither central office personnel nor school administration asked to take an active part in the study, both requested information, by way of a reading list, about action research. Both school administration and central office personnel believe a similar approach could be explored in other areas.

The beginning of the project coincided with the arrival of seven student teachers. Three of the student teachers were placed at Northern and I acted as faculty consultant. All three of the student teachers had gained experience with portfolios as university students. In many ways the student teachers provided an important link between student and teachers, having experienced portfolio assessment both as teachers and students.

Although only two of the eleven science teachers from Northern High School were included in the study by way of interviews, the action research group was open to all members of the science department. All but two teachers attempted some form of alternative assessment during the study, and the two that didn't did not express any negative reactions to the project, they merely wanted more time to examine what was going on in their colleagues classes.

Four consecutive weekly meeting were held at Northern High School to examine alternative assessment, with nine monthly meeting to follow. Each meeting lasted approximately 40 minutes and was attended by the majority of the science department. The student teachers attended the first five meetings. A few teachers, because of coaching responsibilities, attended only a few of the meetings.

Southern High School:

The relatively new school setting at Southern High School and the slightly younger age group of teachers provide many parallels with the Northern High School group. The school population could be classified as middle-class. The school was considered reasonably well-equipped and had an experienced full-time laboratory aide. Little teacher change had occurred in the first eight years of school operation. During the time of the research study, Southern High School was a pilot project looking at Site-Based-Management. Establishing procedures to accommodate decentralized funding became the primary focus for teachers in the school. The school administration expressed no concerns about teacher involvement in alternative assessment and no follow-up information was ever requested.

At Southern high school, the teachers identified as leaders within the science department were also heavily involved in extra-curricular activities, especially coaching. This involvement made organizing action research meetings difficult. Although four weekly meetings were scheduled to begin the project only two took place. Last minute emergency meeting, called by the principal, to examine a new budgeting system, took priority. In addition, regularly scheduled monthly meetings were often rescheduled - only two monthly meeting of the six took place. At Southern High School only two teachers from the seven member department became actively involved in the action research project. Meetings never had more than four people in attendance, and often a different four people at each of the meetings. Meetings tended to take about 30 minutes.

Eastern High School

At the beginning of the research project, Eastern High School had a cohort group of long-serving experienced teachers. During the fall of 1993, three new science teachers joined the department at Eastern High school. Students from Eastern High School tend to come from middle class to lower middle class families. Once an academic school, the population has changed greatly over the past ten years. Ten successive years of a declining student population had been dramatically reversed in 1992. During the action research study, Eastern high school experienced a resurgence in student population. Because the staff did not have to worry about forced transfers linked with a declining enrollment, morale had improved greatly.

The research site presented a special challenge for me. As a new teacher in the school, beginning in February 1993, I was looked upon to provide leadership in developing assessment strategies. Administrative support was very strong in the school.

The project initiated with four weekly meetings followed by 10 monthly meeting for the entire 1993/94 school year. The following school year, monthly meetings continued. Often meetings were organized around a social function and often took place outside of the school. Most meeting took 60 minutes, but it wasn't unusual for two hour meetings if the refreshments were good. By the second year (1994/95), members of the mathematics and social studies departments often joined us for the meetings.

Five of the ten members of the science department were included in the study. All but one member of the science department regularly attended the meetings. All members of the science department at Eastern attempted at least one type of alternative assessment.

Context for the interviews

What follows is an outline of teachers who were interviewed as part of the study. It is important to note that not every teacher who became involved in the project was selected for an interview. The outline is designed to provide the reader with a context for interpreting teacher comments, presented later in the study. Experience, subjects taught, and gender are provided as possible factors that might influence their perceptions about alternative assessment.

An in-depth profile of the interviewees is presented after the outline. The detailed descriptions were largely those of the individuals involved in the study. In all cases, teachers approved of their descriptions, prior to including them in chart form.

Teacher Interviewees

Teacher name	School	Teacher Experience years	Subjects	Member of pilot study
Marvin	Eastern High School	26	Biology chem science 10	no
Henry	Northern High School	20	Chem physics	yes
Albert	Northern High School	17	Biology science 10	yes
Jane	Eastern High School	1	science 10	no
Thad	Eastern High School	25	Chem science 10	no
Mary	Southern High School	13	chem science 10	no
Gary	Southern High School	11	chem science 10	no
Larry	Eastern High School	1	Science 10 Science 14	no
Gordon	Eastern High School	7	Physics science 10	no

In-depth profile of teacher interviewees

Teacher name	Profiles
Marvin	- Spent the last 14 years of his teaching career at Eastern High School. - Self-described as reluctant about changing teaching strategies and modes of assessment. - Was not active in curriculum development or assessment beyond the school level. Had never attended an ATA Science Council Conference prior to becoming involved with the action research group. - Observed alternative assessment strategies for one year in the school before implementing any in his own class. Most reluctant to change assessment strategies at the grade 12 level. - Class structure predominantly lecture-based approach to teaching, although he began using many more laboratories as a member of the action research group. Debates and peer assessment were never used during the study.
Henry	- Recognized leader in curriculum development at the school, system, and provincial levels. - Very energetic and creative teacher who is well respected by students and teachers. - Department head and curriculum leader within the school. (The only science department head the school has had). He became the major driving force behind the action research group at Northern High School. - Developed strong interest in performance-based assessment strategies. A very strong commitment to student-designed laboratory investigations. - Not reluctant to use alternative assessment at any grade level.
Albert	- One of the teachers who helped develop the science area at the Northern High School. - Not reluctant to change teaching strategies and modes of assessment. Willing to try alternative assessment at any grade level. Actively worked to develop new assessment strategies. - Active in curriculum development and assessment at the school level, but not at a provincial level. - Recognized as a strong teacher by students and staff. He had developed the confidence needed to implement changes. - Class structure predominately a lecture-based approach to teaching, although he began using many more laboratories as a member of the action research group.

Jane	• Entire two years of her teaching career at Eastern High School. • Confident, very organized teacher. Recognized as a good teacher by staff and students. Worked with action research group to develop and modify various assessment strategies. • Utilizes cooperative learning strategies, debates, and a variety of laboratory investigations. A strong interest in science-related social issues. • Taught grade 10 and 11 courses. • Science major, with a strong academic background.
Thad	• Spent his entire teaching career at Eastern High School. • Self-described as reluctant about changing teaching strategies and modes of assessment. Focused on providing feedback, rather than working with the action research group to develop new assessment strategies. • Was not active in curriculum development or assessment beyond the school level. Had never attended an ATA Science Conference prior to becoming involved with the action research group. • Observed alternative assessment strategies for one year in the school before implementing any in his own class. Most reluctant to change assessment strategies at the grade 12 level. • Recognized as an organized and dedicated teacher by students and staff. Former department head, who was not reluctant to give up his position in the midst of school reorganizations. • Class structure predominantly lecture-based approach to teaching, although he began using many more laboratories as a member of the action research group. Began using performance-based-assessment after observing the process for 12 months.
Mary	• Spent the last 10 years of her teaching career at the school, the last four as department head. Recognized leader in curriculum development at the school, system, and provincial levels. • Had been involved with various forms of field testing, Alberta Education, developed materials for assessment. Focused action research on providing feedback, rather than working with the action research group to develop new assessment strategies.

	• Very energetic teacher who is respected by students and teachers. Involved in coaching and a variety of school committees. Has served on a number of Alberta Education committees for curriculum and evaluation. Continues to do a great deal of work with Student Evaluation. • A recognized curriculum leader within the school. She initiated the action research group at Southern High School. • Developed strong interest in performance-based assessment strategies. A number of concerns about using portfolios where students could choose some of the work to be assessed. • Cautious about changing assessment strategies. Spoke often of accountability to parents and the school administration. Most reluctant to use alternative assessment at grade 12.
Gary	• Spent the last 10 years of his teaching career at Southern High School. Has recently began to do work in curriculum at a provincial level. • Very energetic teacher who is respected by students and teachers. Involved in coaching. • Had been involved with various forms of field testing, Alberta Education, developed materials for assessment. • A number of concerns about using portfolios because students might not receive feedback from teachers until well into the school year; however, he liked using journals. • Cautious about changing assessment strategies. Spoke often of accountability to parents and the school administration. Most reluctant to use alternative assessment at grade 12.
Larry	• First year of his teaching career at Eastern High School. A Social Studies major and science minor, who had been hired to teach primarily in the science area. • Confident teacher who is well liked by students. Larry looked for assistance, especially from Jane, during his first year. • He was re-assigned to a junior high school following the 1993/94 school year. Was a part of the action research group for one year. • Taught grade 10 and 11 courses.

Gordon	• First year teaching at Eastern High School 1993/94. He replaced a teacher on leave. Gordon was given a position at another high school the following year. A member of the action research group for one year. • Prior to arriving at Eastern High School, he did his previous teaching in rural Alberta at the junior high school level. • Gordon was a risk-taker. He readily attempted new assessment strategies that he had personalized at all grade levels. • Included in the study for one year. An energetic teacher that became very interested in computer assisted instructions.

Significance of the study

The use of performance-based laboratory assessment, portfolios, and student journals that emphasize metacognition was an attempt to provide a wider view of science and better reflect the synchrony of curriculum intent, learning activities, and student assessment. Those who have leapt to portfolio assessment, journals, or performance-based-laboratory assessment fearing that they would miss the "approach-of-the-day", without considering practical problems such as storage, criteria for selection, or anticipating pedagogical problems have often been disappointed by the results (Maeroff, 1991). Teachers must first decide: what should be assessed; who decides what is assessed; and how these elements are to be assessed. Failure to do so has caused aspects of alternative assessment to implode upon themselves (Wiggins, 1989). Pedagogic implications and practical limitations must be addressed. Most importantly, strategies that increase the amount of time spent on evaluation at the expense of curriculum planning are not always deemed beneficial by either students or teachers (Wolf, Bixby, Glenn, & Gardner, 1991). The following objectives were considered when establishing action research groups.

For new assessment strategies to work we must:

- seek to develop assessment strategies at the school level that take into account such practical problems as storage space, access to computers and electronic support, diversity of teaching assignments, and the number of students being assessed (O'Neil, 1992; Wiggins, 1992b).
- be sensitive to increasing the amount of time teachers spend assessing students. Ideally, new assessment strategies will shift the emphasis from teacher as "marker" to teacher as "facilitator and adviser", allowing teachers to devote more time for curriculum development and

less time for marking and associated clerical tasks (Eisner, 1993; Rogers, 1992; Maeroff, 1991 Madaus & Kellaghan, 1989).

- encourage students to take increased responsibility for participating in their own assessment (Maeroff, 1991; Madaus & Kellaghan, 1989; Eisner, 1985). The new assessment strategies must shift the focus of assessment from how the teacher values assignments to the value that the assignments provide for the student (Eisner, 1993; Madaus & Kellaghan, 1989).
- provide teachers and students time to reflect upon assessment strategies and evaluate their effectiveness (Tierney, 1991; Tobin, et al., 1988; Vavrus, 1990; White, 1988).

New strategies for assessment provide exciting opportunities, not hard-and-fast solutions to old problems. Therefore, the action research project was much more than collective problem-solving. Documenting the successes and failures of the school-based action research groups, as we worked to implement assessment strategies that better reflect learning, has provided important signposts to others interested in similar problems. Perhaps the most significant part of the study is the focus on student perceptions of assessment strategies that were designed to provide more democratic classrooms.

Delimitation

The study was conducted in three urban high schools. The teachers chosen for interviews were those who had made a commitment to change assessment practices. These teachers indicated dissatisfaction with the perceived division between current assessment strategies and curriculum intents. No attempt was made to restrict the size of the action research group in any of the schools; however, not every teacher was interviewed. Because I requested interviews from teachers who held strong opinions, no claim of a random sample will be made. I leave it to the reader to interpret whether the views expressed by teachers during the interview reflect the beliefs and opinions of other teachers. Gadamer explains:

> When we try to understand a text, we do not try to transpose ourselves into the author's mind, but if one wants to use this terminology, we try to transpose ourselves into the perspective within which he has formed his view. But this simply means that we try to understand how, what he is saying, could be right. We try to make the argument stronger. (1993, p. 292)

All students chosen for interviews came from my classes. Being selected for the interviews had an unexpected effect, students felt special because I considered their opinions important. A few students not selected for interviews asked why they weren't chosen. Although colleagues from all research sites actively sought student feedback, no other member of the research group used taped interviews. Student reactions from one class are not assumed to represent reactions held by all students in my class, nor do I assume that they reflect a cross section of opinions that students would hold in other classes. Each class has a unique character. Teacher influence must also be considered.

The action research study was not designed to provide a step-by-step plan to initiate change. However, it explored how the teachers, in each of the schools, attempted to bring about changes in assessment strategies. Each strategy was viewed within the context of the school community and teacher personal knowledge. Eisner (1985) indicates that claims of validity can be defended by providing descriptive evidence. Validity can only be judged by the reader. Reliability, as traditionally denoted in terms of repeatability, will not be claimed. No attempts were made to repeat an identical study in each of the three schools. Nor was there an attempt to replicate a set of procedures in each classroom regardless of teacher interests or beliefs. Factors identified as temporal, spatial, and ecological within each of the schools and classrooms are too diverse to replicate.

Limitations

I believe that a framework of trust established prior to the interview is crucial. Critiques of assessment strategies have greater potential to be genuine if both interviewee and interviewer share a collective concern. Interviews were only conducted once a rapport was established. It should be noted that teacher and student opinions changed dramatically, as the action research project progressed. A difficult day of teaching or a disappointing mark on an assignment amplifies the problems and concerns. Equally important to acknowledge is that "good days" yield positive interviews. Interviews represent "snap shots" in the experiences of teachers and students who worked to seek more authentic assessment strategies. A complete record of teacher thinking was requested by way of journals, but abandoned early into the program when I discovered that they were not being done. Adding teacher journals to a research project, designed to reduce the amount of time teachers spend at clerical tasks seemed counter-productive.

Difficulties of acting as both teacher and researcher have been well documented (Tierney, 1991; Tobin, et al., 1988). Claims of objective-observer will not be made.

Because student interviews were elicited from my class alone, the findings are not transferable to other classes. Student responses are highly contextualized and affected by a number of factors, including their relationship with the teacher, their perceived importance in the class, the cohesiveness of the class, and expectations placed upon them and accepted by them. Student responses can only be interpreted as an indicator of what might occur in a classroom. They also serve as a reminder to teachers that all students might not hold an opinion that teachers have come to accept as being representative. The work of Berg & Brouwer (1991) indicates that teacher perceptions of students' ideas are not always accurate.

Assumptions

The research is premised on the belief that student assessment in science can be improved by employing a wider array of strategies that enable students to become more actively involved in their own assessment. The idea of improvement links the manner in which students learn, the manner in which curriculum is encountered and the manner in which students are assessed. Improvement of assessment is not defined as improving test scores for standardized tests or improving the correlation between teacher assigned grades and those achieved from standardized testing. The meaning of improvement is linked with the notion that the assessment strategies are more meaningful for both teacher and student (Aoki, 1984).

The first assumption is based upon two more fundamental assumptions, that students learn in different ways (Gardner, 1993) and that because current assessment strategies do not take into account differences the true abilities of many students are not recognized (Wiggins, 1993).

Definition of terms

Action research: Action research is a form of collective self-inquiry undertaken by participants to improve the rationality and justice of their own social or educational practice. In addition, the inquiry attempts to provide understandings of educational practices (Kemmis, 1990). To be action research the approach must be collaborative, though it is important to realize that the action research of the group must be achieved by critically examining the actions of individual group members. The critical feature highlights the terms *action* and *research*, which means: trying out the ideas in practice as a means of improving the curriculum. Action research links theory with practice.

Assessment: The root of the word assessment comes from the Greek, meaning to "sit with" (Maeroff, 1991). The origin suggests assessment was subject focused rather than group norm referenced. Grant Wiggins (1992a) indicates that focus on the individual enables teachers to determine if the child is saying what he/she really intends to say.

Authentic assessment: The testing situation parallels the learning situation. Unlike traditional testing situations, students would be allowed to ask for clarification of the question or the task in authentic assessment. Assessment practice would not be restricted to pencil and paper format. Computer data bases, information systems, and research skills would be integrated within the assessment approach that would have multiple boundaries of time, location, and facilitated assistance, depending upon the context of the task. Ryan (1994) describes alternative assessment as the process of gathering evidence and documenting student growth in a meaningful context.

Evaluation: The process of making carefully determined value judgments and issues (Doran, Lawrencz, & Helgeson, 1994). Smith (1988) links evaluation to its historical roots of "value" and more distantly to "valor", meaning derived from power. He explains that valor is more than bravery. By placing value on something, you are obligated to act. Eisner (1993) uses a similar notion of evaluation as a one-way power relationship between teachers on students. The teacher imposes his or her values on students.

Metacognition: For this study metacognition will be referred to as thinking about thinking; being aware of one's own thinking.

Measurement: According to Doran et. al. "measurement has generally been defined as the process of testing, but it should also include a more encompassing term that has included the use of observations, discussions, as well as pencil-and-paper tests" (1994, p. 299).

Middle language: Refers to the language of compromise that develops through discourse. Desmond (1987) refers to the language of self and other as "metaxological" or discourse in the middle. Unlike the dialectical which describes self as different from other, the middle language begins to describe self in terms of interactions that change both self and other.

Performance-based assessment: Refers to assessment tasks that more closely mirror the tasks that students accomplish while learning. For science students, process and psychomotor skills, commonly associated with investigative laboratory work, are focused upon (O'Neil, 1992). The goal of performance assessment has been to develop activities that permit students to pursue an experimental inquiry focusing on process skills and construction of new knowledge (Shavelson, 1992).

Portfolios: Represent collections of student work. Vavrus indicates that the portfolio is more than "just a container full of stuff" (1990, p. 48). She explains that the portfolio is a systematic organized collection of evidence used by the teacher and student to demonstrate the student's growth of knowledge, skills and attitudes. Ryan (1994) describes portfolios as collaborative work done by student and teacher in selecting what work is to be assessed.

Science-technology-society: An STS approach to science attempts to make the learning of science personally relevant and socially meaningful (Aikenhead, 1980). Scientific knowledge is linked with technological innovation, the manner in which students experience science, in order to close the gap between classroom study and life experience (Ritter, 1991). By placing science within a social context, it is hoped that students gain a better appreciation for the impact of science and technology on society. Equally important is a view of science as a social activity (Bronowski, 1977). Science and the construction of scientific knowledge are greatly affected by society.

Chapter 2: The Rationalist Model for Assessment

Overview

In this chapter I will provide the philosophical underpinnings for developing an understanding of why current assessment strategies need changing. In doing so, I will trace the rise of rationalism as an alternative religious doctrine. The chapter begins by looking at why Cartesian rationalism became the dominant way of thinking and of valuing knowledge. The underpinnings of rationalism are examined to identify assumptions inherent within this philosophical orientation. These often taken-for-granted assumptions dictate what is to be valued by limiting the types of questions that can be investigated. My assertion is that rationalism is valued because it provides a methodology for problem-solving that ensures access to power. Later in the chapter, the humanist alternative to rationalism is presented. The purpose of this perspective is not in replacing rationalism, but in providing another window to view and interpret natural events.

In the second part of the chapter the limitations of rationalism are explored within an educational context as it applies to student assessment. The acceptance and widespread use of the standardized assessment models for assessment are explained by linking them with rationalism. What is defined as important is that which can be counted. Assessment practices that extend beyond the confines of rationalism, by employing expert judgment or connoisseurship, provide an alternative. Authentic assessment practices are those which do not attempt to separate the subject from the knowledge to avoid bias, but investigate the journey of inquiry within the context in which learning occurred.

In the third part of the chapter, I have documented obstacles to change and provided an outline of previous research in Britain, the forerunner of alternative assessment. Much can be learned by exploring problems arising from previous experiences.

Influence of the Cartesian rationalism

According to Plato, Socrates assigned citizens of the Republic to three classes: rulers, auxiliaries, and craftsmen. Socrates believed a stable society demanded that intelligence be used to identify the leaders. In sharp contrast to the world Socrates lived in where birth right provided access to power, his system for the selection of leaders would be based upon merit. Plato dreamed of a world ruled by philosopher kings, who used mathematics, the purest form of knowledge for reasoning in decision-making (Lindberg, 1992). According to Plato, all that existed in the world around him was an imperfect reflection of reality. Reality could only be understood by mathematics, because it is founded on abstract reasoning, unencumbered by an imperfect physical world. The abstraction of reason from the physical world led Enlightenment philosophers, such as Descartes, to the assumption of dualism that the mind is separate from the body. The dichotomy of mind and body became so well entrenched in Western Philosophy that it was often accepted as the axiom or self-evident truth on which other ideas were formulated (Dimasio, 1994a).

The ancient Greeks identified reason *(logos)* as a virtue. This identification ensured that reason became the highest form of knowing. Being a virtue, it was assumed that reason would show the way to truth. The Romans reinforced the notion. During the middle ages, the Catholic Church and theologians narrowed the notion to reinforce revealed truths. Rational action leads to good (Lindberg, 1992).

The "Age of Reason" has been with us for nearly 500 years. Although the definitions for reason, like so many definitions, have changed greatly over time, the notion that reason legitimizes truth has been long enduring. While philosophers kept busy arguing notions of what constitutes reason, social leaders embraced its power to validate their actions. What is most important are not the definitions of philosophers, or tracing the evolution of meanings but understanding what society believes it to be. Saul (1992) indicates that what is most important is our expectations of "reason" and the mythology surrounding the word.

Historically, the Age of Reason provided an alternative to what many considered to be the arbitrary power of church and state (Groome, 1980). Descartes' deductive inquiry and abstract arguments drew acceptance because they were supported by mathematical relationships. In that the axioms of mathematics were shown to be immutable, the acceptance of reasoning as a harbinger of truth was assured. Reasoning produces answers that are unchangeable. Any truth, if reasonable on Tuesday, should also be true on Wednesday. In addition, because reasoning is systematic, it can be repeated. Repeatability permits predictions. Prediction serves a pragmatic purpose, assuring widespread usage. Most importantly, the new paradigm supported scientific investigation and provided an alternative to theology as the dominant way of viewing the world (Dimasio, 1994a; Lindberg, 1992). As science replaced religion as the dominant manner in which knowledge was collected, reason assumed ascendancy as a way of knowing. So dominant was reason that rationalism quickly distanced itself from other human characteristics: spirit, appetite, emotion, will and most importantly experience (Saul, 1992).

Polanyi in *Personal Knowledge*, explains how qualities other than reason supported by deductive logic were considered untrustworthy. He interprets Descartes by writing:

> Descartes had declared that universal doubt should purge his mind of all opinions held merely on trust and open it to knowledge grounded on reason. (Polanyi, 1962, p. 269)

When the scientific way of viewing the world supplanted a world view supported by theology, reason became the arbiter and purveyor of truth. The certainty of knowledge exposed by priests or kings had been replaced by those who used reason. To gain acceptance, other forms of knowledge were required to provide solutions that did not conflict with those acquired by reason. Gadamer explains why reason has usurped tradition as arbiter of truth:

> In general, the Enlightenment tends to accept no authority and to decide everything before the judgment seat of logic. The written tradition of Scripture, like any other historical document, can claim no absolute validity; the possible truth of the tradition depends on the credibility that reason accords. It is not tradition but reason that constitutes the ultimate source of authority (1993, p. 272).

Conclusions supported by logic were referred to as "rational". Supportive evidence came to be known as "facts". The systematic approach of rational inquiry allowed many different thinkers to find the same conclusion by assembling the "facts", and hence, the fact took on the meaning of immutable truth. Irrational views, those not founded upon reason, were deemed less valued. Views linked with emotion, or experience, were deemed untrustworthy if they were not supported by reason. Eventually, the idea that irrational thought might cloud rational thought and disguise that which is true emerged. Not only was irrational thought considered inferior to rational thinking, it was considered a deterrent, something to be guarded against (Dimasio, 1994b).

The notion that logic should not be invaded by emotion is well-documented (Dimasio, 1994a). Plato and Kant caution against mixing logic with emotion. Perhaps the strongest expression of caution was issued by Descartes, who celebrated the separation of reason from emotion and mind from body. However, despite its long-held appeal, many modern-day philosophers provide cautions against the use of reason to the exclusion of other ways of knowing (Bronowski, 1977; Medawar, 1983; Polanyi, 1962; Saul, 1992). Dimasio (1994b) argues that the greatest questions that face our society today are not those that can be solved by reason but by emotion. Drawing from his studies of patients, who have had personality changes due to brain disorders or injuries, Dimasio concludes that the absence of emotion is problematic. Sociopaths, for example, are capable of many logical operations, yet they lack an ability to interact with others in society. He states:

> It does not seem true that reason stands to gain by operating without the leverage of emotion (Dimasio, 1994b) p. 144).

Timpane points out the folly of abandoning emotion from the roots of knowledge. He states:

> At the end of every successful argument, no matter how weighty the evidence or powerful the reasoning, the beholder performs a non rational act: the leap of acceptance. It may be short - one may feel pushed — but a leap it always is (Timpane, 1995, p.104).

Notion of the superiority of objective knowledge

Cartesian dualism separated mind and body, in an attempt to elevate reason and subordinate other characteristics of knowing. By suggesting that reason employs objective knowledge, Descarte proposed a methodology for finding truth. Objective knowledge, that which comes from reasoning, was equated with truth; while subjective knowledge, that which comes from emotion and experience, was considered of lesser value. Nowhere has the bond between objective knowledge and truth been so strongly forged than in the natural sciences. Empirical evidence, that which was directly observed, was considered to be removed from human interference (Lakatos, 1974). Polyani (1962) describes the dangers of abandoning belief, based upon experience, to the exclusive use of what is believed to be objective knowledge:

> Belief was so thoroughly discredited that, apart from special privileged opportunities... modern man lost his capacity to accept any explicit statement as his own belief. All belief was reduced to the status of subjectivity; to that of an imperfection by which knowledge fell short of universality.
>
> We must now recognize belief once more as the source of all knowledge. Tacit assent and intellectual passions , the sharing of an idiom and of cultural heritage, affiliation to a like-minded community; as such are the impulses which shape our vision of the nature of the things on which we rely for our mastery of things. No intelligence, however critical or original, can operate outside such a fiduciary responsibility. (Polanyi, 1962, p.266)

Polyani (1962) refutes the idea of objective knowledge and replaces the idea of objectivity with Personal Knowledge.

> Key objectivity, as it applies to the exact sciences is a delusion and is in fact a false ideal. (Polanyi, 1962, p. 22)

Quantification, the by-product of rationalism, has created an even greater reduction. Following the Platonic tradition, mathematical algorithms are summoned for support, and numbers, because they do not reside within the body, are described as objective. Reason, therefore, by way of numbers, stands for truth. However, the real danger of rationalism is the idea that reason, if done correctly, is fool-proof.

> Enlightenment, namely that methodologically disciplined use of reason, is unable to safeguard us from all error. This was Descartes' idea of method. (Gadamer, 1993, p. 277)

Embedded within the teleology of rationalism is the notion that numbers do not make value judgments and that empirical evidence provides the numbers. Interpretations performed to acquire the numbers are not even considered. Rather than interpret natural events, scientists are portrayed as people who find the numbers which allow us to measure reality. The number, reified in such a manner, becomes interchangeable with reality. Inductive logic allows for generalizations which become the theories. In turn, the theories enable predications, which confirm the notion that reality is "out-there" just waiting to be discovered. The acceptance of this orientation enables humans to exert control over the world. A world-view of "man and the world" directs actions.

Subjective knowledge, that which derives from experience, emotion, or intuition was deemed untrustworthy. This knowledge, because it resides within the subject is linked with bias or prejudice. Gadamer indicates:

> The history of ideas shows that not until the Enlightenment does the concept of prejudice acquire the negative connotations familiar today. Actually prejudice means a judgment that is rendered before all elements that determine a situation have been fully examined. (Gadamer, 1993, p.270)

Following a rationalist view, numbers, because they lack bias or prejudice, now stand for truth. Gould, a scientist, also denounces the preoccupation with quantification. Numbers, argues Gould (1981), are often used as objective purveyors of truth, yet history tells us that the numbers are often employed to support *a priori* conclusions. In critiquing those who believe that numbers by themselves demonstrate truth, he states:

> Numbers suggest constraint, and can refute; they do not by themselves, specify the content. Theories are built upon interpretations of numbers, and interpreters are often trapped by their own rhetoric. They believe in their own objectivity, and fail to discern the prejudice that leads then to one interpretation among many consistent with their numbers. (Gould, 1981, p. 76)

Gould (1981) refutes the idea that scientific knowledge operates on reason alone. He explains that science progresses by hunch, vision, and intuition; all forms of what has been deemed subjective knowledge. Theories are not inexorable inductions from facts. The most creative theories are often imaginative visions imposed from facts; the source of imagination is also strongly cultural. Gould explains:

> But science's potential as an instrument for identifying the cultural constraints upon it can not be fully realized until science gives up the twin myths of objectivity and the inexorable march toward truth. (Gould, 1981, p. 22)

The idea that objective knowledge removes bias has also been challenged (Lakatos, 1963; Kuhn, 1970; Feyerbend, 1975; Bronowski, 1977; Medawar, 1983; Dennett, 1991). Prejudice also resides within the rationalist approach. Rationalism, in an attempt to provide a systematic method ensuring a pathway to truth, creates bias. It restricts the type of questions that can be asked. Those that can not be solved by reason alone are often considered less important.

> Bias in one's observations and conclusions is a function of not what is put in but what is inadvertently left out. If we leave out of our considerations and reports what our vehicle cannot contain, we provide a limited, indeed a distorted perspective. (Eisner, 1985, p. 7)

Gadamer explains that any movement to overcome prejudice is limiting. By focusing solely on eliminating prejudice, a barrier is placed around the knowledge which we create and value. He states:

> The overcoming of all prejudice, this global demand of Enlightenment, will itself prove to be a prejudice, and removing it opens the way to an appropriate understanding of the finitude which dominates not only humanity but also our historical consciousness. (Gadamer, 1993, p. 276)

Gadamer indicates that the negative notion of prejudice has been linked with the fact that pre-judgment lacks a methodology. He explains that:

> The only thing that gives a judgment dignity is its having a basis, a methodological justification (and not the fact that it may be actually correct). For the Enlightenment the absence of such a basis does not mean that there might be another kind of certainty, but rather that the judgment has no foundation in things themselves. (Gadamer, 1993, p. 271)

Polanyi also opposes the notion that reason illuminates the way to a single truth by removing bias. He states:

> For once men have been made to realize the crippling mutilations imposed by an objectivist framework - once the veil of ambiguities covering up these mutilations has been definitively dissolved - many fresh minds will turn to the task of reinterpreting the world as it is, and then once more will be seen to be. (Polanyi, 1962, p. 381)

Method and power

Saul (1992) examines why reason, under the guise of objective knowledge, has systematically devalued other human characteristics. He concludes that subjective knowledge, acquired by emotion, intuition, or experience, provides no learnable and repeatable system that ensures truth. Logic, by contrast, provides a formula for determining that which is true. Saul indicates:

> We carefully – rationally in fact – assign blame for our crimes to irrational impulse. In this way we merely shut our eyes to the central and fundamental misunderstanding; reason is no more than structure. And structure is most easily controlled by those who feel themselves to be free of cumbersome weight represented by common sense and humanism. (Saul, 1992, p. 16)

A link between rationalism and power is forged when those who apply logic to formulate answers are deemed experts. The relationship between experts and power can be found in the word *"technocrat".*

> Technology is a relatively new word, combining the Greek *'techne'* (skill) with *'logos'* (knowledge). The skill of knowledge. But the noun technocrat has a very different meaning. *Techne* in this case is attached to *'kratos'* (strength, power). The technocrat's skill lies in his exercise of power. The skill of power. His is an abstract profession involving only narrow bands of knowledge. (Saul, 1992,. p. 107)

A method derived from reason and logic provides answers. The notion of "expert" is supported by clearly defined and reproducible answers. Experts without definitive answers are held with skepticism. What good is an expert who can not provide you with answers? The greater the certainty projected, the greater is the value and influence of the expert.

Certainty can even be used to define the importance of a question. Questions that do not allow themselves to be measured are often relegated to a secondary position, that of opinion. Certainty, assuming that it is supported by measurement to make it exact, can even be used to stratify experts according to their access to power. When two experts provide different answers, one is often considered correct at the expense of the other. Those with the best numbers are often most respected. At best, both might be rejected while acceptance or rejection awaits further analysis. However, such reductionism blurs the importance of the subject as knowledge is constructed and the dynamics of language. Knowledge does not exist isolated from the subject waiting to be found. Smith (1988) points out the expectation for an unambiguous, single answer subsumed in the Enlightenment tradition. He states:

> An orientation to technology inspired by the rise of science creates a demand for precise and univocality in language, which in turn relies on a definitional approach to meaning rather than a discursive one. Words are defined to mean one thing and one thing only, a condition predicates computer languages and word processing / information science industries (Smith, 1988, p. 227)

However, language, by its very nature, is ambiguous (Langer, 1957; Levine, 1985). Ambiguity creates discourse and prevents a single pathway to an immutable and ever present truth. The rationalist approach, supported by inductive and deductive reasoning, designates a single pathway to truth, thereby limiting our understanding (Smith, 1988). Gadamer indicates that language serves an important vehicle in opening up questions rather than drawing single conclusions. He explains:

> The essence of the question is to open up possibilities and keep them open. (Gadamer, 1993, p. 299)

Looking away from single answers to consider the multitude of possibilities, as suggested by Gadamer, requires a very different philosophical orientation. What we need is a humanist orientation that recognizes that emotion (Dimasio, 1994a), personal knowledge (Polanyi, 1962), common sense (Bronowski, 1981), and virtue (Saul, 1992) do not provide univocal answers. Univocality narrows the discussion, limits possibilities and confines the type of questions that can be asked.

> The secret, then, is that we must alter our civilization from one which feels satisfaction, and not anxiety, when doubt is established . (Saul, 1992, p. 584)

Ignoring doubt to embrace certainty will not provide truth; however, certainty does provide access to power. The linking of rationalism with power has ensured it long standing existence, despite its many failures and shortcomings. To abandon rationalism is to abdicate authority.

> The essence of rational leadership is control. To admit failure is to admit loss of control. (Saul, 1992, p.11)

Saul explains:

> The technocrats suffer from the character defects which have to do with their inability to maintain any links between reason, common sense, and morality. They believe themselves to be the inheritors of the Age of Reason, and therefore, do not understand why their talents fail to produce intended results. Their abstract view of the machinery of human society prevents them from understanding the natural flow of events and from remembering when they themselves have erred and why. (Saul, 1992, p. 107)

To arrogate method

So dominant is the rationalist methodology, as a harbinger of truth, that it has been adopted by the social sciences. Gadamer explains:

> The logical self-reflection that accompanied the development of the human sciences in the nineteenth century is wholly governed by the model of the natural sciences. (Gadamer, 1993, p. 3)

Saul (1992)) explains that the human studies were once ridiculed because they did not conform to the methodology supported by rationalism. There was little choice but to conform. He clarifies that:

> Not only have the humanities been singled out as the enemy of reason, but there has been a serious attempt to co-opt them by transforming them into a science. (Saul, 1992, p. 131)

The exclusive use of a scientific orientation for human studies narrowed the focus of what was studied. Only questions that could be answered by reason became worthwhile knowing. Relationships supported by measurement were most valued. Human studies founded upon reason became known as human sciences. Gadamer explains:

> Human sciences is too concerned with establishing similarities, regularities, and conformities to law which would make it possible to predict individual phenomena and processes. (Gadamer, 1993, p.4)

The methods of science are based upon finding results that are confirmed by repeatability. If human events and the knowledge derived from those events are validated by duplication or imitation, only a small spectrum of human behavior can be understood. Truth is not measured by the numbered of times an event is observed. Gadamer explains:

> But the specific problem that the human sciences presents to thought is that one has not rightly grasped their nature if one measures them by a yardstick of a progressive knowledge of regularity. (Gadamer, 1993, p.4)

Eisner (1985) echoes Gadamer's assertion within an educational context. He explains:

> Scientific activity yields propositions so that truth can be determined in relation to its instrumental value, a value dependent upon its predictive or explanatory accuracy. Artistic activity creates symbolic forms which themselves present directly an idea, image, or feeling which resides within rather than outside of the symbol. (Eisner, 1985, p. 89)

The movement to adopt a scientific model for human studies is founded upon a distorted view of the natural sciences. Scientific facts, presented as an irrepressible march toward truth, have been severely criticized (Bronowski, 1977; Feyerbend, 1975; Lakatos & Musgrove, 1974; Kuhn, 1970; Lakatos, 1963). Equally unpalatable is the idea that facts produce truth (Medawar, 1983; Bronowski, 1981; Lakatos & Musrgove, 1974). In rejecting the link between truth and a collection of facts, Saul explains:

> The rational revolution has left us with the conviction that truth is a fact or a compendium of facts. This has grown into a way of life, which now turns on to structures and expertise. Most individuals with some expertise or authority work within these structures and therefore have control over an element of modern truth. (Saul, 1992, p. 281)

The idea that objective knowledge alone can act as the gauge to affix value completes the circle of reasoning created by rational thinking. Polanyi explains:

> For modern man has set up the ideal of knowledge, the conception of natural science as a set of statements which is 'objective', even while its presentation may be shaped by convention. This conception, stemming from a craving rooted in the very depths of culture, would be shattered if the intuition of rationality in nature had been acknowledged as a justifiable and indeed essential part of scientific theory. That is why scientific theory is represented as a mere economical description of facts; or as embodying a conventional policy for drawing empirical inferences; or as a working hypothesis, suited to man's practical convenience - interpretations that all deliberately overlook the rational core of science. (1962, p. 16)

The tradition of dualism, that separates the subject from the object, only distorts knowledge. Knowledge, because it is created within the subject, is both subjective and objective. The historical preoccupations to look for objective knowledge apart from a subject exacerbates inquiry in human sciences.

> ... historical research is carried along by the movement of life itself and cannot be understood teleologically in terms of the object into which it is inquiring. Such an "object" itself does not exist. This is precisely what distinguishes the human sciences from the natural sciences. (Gadamer, 1993, p. 285)

To understand is to acknowledge both subject and object.

"Full understanding can only take place within this objective and subjective whole." (Gadamer, 1993, p. 289)

Rationalism as a method for education

It should not be surprising that rational thinking has become the dominant orientation for education. Educational research has a tradition of borrowing techniques from the sciences and social sciences (Eisner, 1985).

> The task of educational research was to treat educational practice as a nomothetic activity, one controlled by laws rather than an ideographic activity, one which was guided by unique characteristics of the particular situation. (Eisner, 1985, p. 88)

He goes on to explain that the:

> ... scientific and technological approaches to schooling lead, . . . , to the attempt to objectify knowledge. Objectivity almost always requires two conditions be met. First, the qualities to which one attends must be empirically manifest, and second, they must be convertible to quantity. (Eisner, 1985) p. 89)

Because objectivity is seen as removing the subject and all of the prejudices inherent within the subject, the rationalist method claims to remove bias. A method based solely on reason, unencumbered by emotion, beliefs, and intuition, was believed to ensure the truth. Eisner critiques the rationalist method that links objectivity with truth. He states:

> When rules are codified and explicit as they are in the three R's ambiguity can be reduced, precision can be increased, and the security of knowing when one is right or wrong obtained. If the syntax of a form of representation is highly rule-governed, it makes it possible to reduce idiosyncratic interpretations, if not eliminate them altogether. (Eisner, 1985, p. 168)

The true allure of the scientific approach is best understood as an access to power. Methodologies that exclusively employ logic, reduce their focus to events that are quantifiable and repeatable and gains credulity from matching predictions with observed outcomes. The greater the match, the greater the certainty, and the greater is the access to power. Eisner explains:

> This yearning for prediction through control was, of course reflected in the desire to make schools more efficient and presumably more effective. Educational research was to discover the laws of learning that would replace intuition and artistry with knowledge and prescribed method. The hunt was on for the best method... (Eisner, 1985, p. 87)

The movement toward greater accountability is fueled by the rationalist method that seeks to quantify the number of matches between prediction and outcomes. Tests, as indicators of future student success, attempt to both define and measure success as a single number. "Good tests" are often defined as good indicators. However, quantitative measurements often fail to tell the whole story: Eisner explains:

> As a result of the partial views that such methods provide, a bias, even distorted picture of reality that we are attempting to understand can occur. In some respects this result is paradoxical because the stringent cannons of the social scientific methodology are the product of a desire to reduce bias and diminish distortion; the claim that they may in fact contribute to the bias and distortion is a severe critique, if true, and a paradox of their intention. (Eisner, 1985, p. 147)

Limitations of the standardized testing model

Standardized testing, the dominant model for student assessment, continues to focus on knowledge components for science courses. Wiggins (1993) explains that standardized assessment sacrifices validity for reliability and is incapable of supporting true curricular change. The traditional approach of standardized testing, by virtue of its intent, is predicated upon providing predictable results. The meaning of one test is defined in relation to how students score on others. Testing, within this archetype, is judged by predictability. The image of a 70% student takes form when the student consistently scores near 70%. Correlations between different exams and even between similar subjects provide predictable indicators that are used to rate the accuracy of individual exams.

Eisner explains:

> For the evaluation of educational practice and its consequences, the single test score is used to symbolize a universe of particulars, in spite of the fact that the number symbol itself possess no inherent quality that expresses the quality of the particular that it is intended to represent. (Eisner, 1985, p. 89)

The reification of a 70% student requires the testing of indisputable facts. Ideally, the facts should be presented in an insular fashion to remove ambiguity. A single, correct answer reinforces predictably and is easily measured. The single correct answer also provides a linear connection between problems and solutions. Working within the confines of this approach, every problem is complemented by a matching solution and the work of the scientist or science student becomes finding the correct fit. Because such linear pathways are readily repeated, they gain acceptance. In an ever-cycling syllogism, single answers are sought because they provide direct cause-and-effect relationships, which in turn are valued because they are repeatable. Repeatability gains acceptance because it provides predictability, something that may only be achieved when a single answer is accepted. Once the assumption that repeatability is the standard for valuing assessment practices is accepted, the circle is completed and any escape becomes impossible. Eisner (1985) cautions against the use of numbers to create meaning. He explains that meaning associated with numbers is derived from two sources: the scale of which they are a part and the referents they are used to represent. To understand what a number means, you must understand its place on the scale and the qualities it is selected to represent. Eisner states:

> No number looks like a referent; numbers are conventions, and the transformation of qualities experienced into such convention *never*, without the ability to imagine the qualities of its referent, can 'contain' those qualities. This means that numbers as surrogates are not self explanatory. They fail to portray by convention the operations that lead to certain scores are replicable and, hence, regarded as objective. (Eisner, 1985, p. 224)

Not surprisingly, the model afforded by standardized testing greatly distorts the nature of science and opposes the Science-Technology-Society (STS) approach to teaching and learning science. Despite learning science within a social context that unveils the ambiguities and doubt associated with making choices, where many times no clear right and wrong exist, assessment strategies focus on single variables and a single correct answer. How can teachers evaluate the effectiveness of science programs if assessment strategies continue to focus exclusively upon decontextualized factual information?

The inadequacy of traditional approaches for contemporary curricula goals is described in detail (Wolf, Bixby, Glenn, & Gardner, 1991). The success of new teaching strategies relies, in part, upon developing assessment strategies that reflect the wider range of goals identified by the STS approach. Tobin points out that:

> There is little reward for changing teaching so as to emphasize high-level cognitive learning and laboratory activities if the assessment systems continue to promote the recall of facts. (Tobin, et al., 1988, p. 451)

The distortion of repeatability and certainty

Once predictability and reliability are embraced as the twin pillars of truth, a methodology is derived to eliminate ambiguity. Certainty is held as an ideal. At first glance, the logic seems impeccable. Given predefined objectives or goals, the student will perform accordingly. The more times similar results are observed, the stronger the truth becomes. Assessment techniques, derived within this paradigm, provide procedure or rules that ensure commonplace standards are established for the testing. The idea that every student is treated in the same manner in every situation ensures predictability and links it with the notion of objectivity.

The rationalist perspective uses objectivity as a compass to the truth. The method in providing certainty guarantees truth. However, the bridge between objectivity and truth creates insurmountable problems when we attempt to evaluate student learning. Because objectivity is seen as the primary virtue of evaluation, qualities that are difficult to measure are often ignored or marginalized. Eisner explains:

> Because a rationalist orientation to procedure provides the baseline criterion that method must meet, those aspects of educational life that are most easily susceptible to measurement command the attention of evaluators. (Eisner, 1985, p. 225)

In describing the dominant approach to assessment Eisner point out:

> ... the ethos of this historical period embraces a systematic and objectified approach to knowledge and necessitates the use of methods that yield conclusions that are replicable. Because quantification is a paradigm for conventionalized descriptions, it is regarded as a necessary condition for achieving objectivity. The ethos of time supports an epistemology that tends to neglect the idiosyncratic and those aspects of educational life that are difficult to objectify through measurement. What results is a biased assessment of the very life we are trying to understand and improve. (Eisner, 1985, p. 225)

To begin looking at assessment differently we must challenge the notion that predictability and repeatability are desirable. Maturana and Varela (1987) explain that true knowledge about knowledge requires an acceptance of uncertainty. They explain:

> The knowledge of knowledge compels. It compels us to adopt vigilance against the temptations of certainty. It compels us to recognize that certainty is not a proof of truth. (Maturana & Varela, 1987, p. 248)

Consider the consequences of assessing student learning exclusively from an instrument that demands certainty. Would we only present things in biology classrooms that had clearly defined answers, knowing that only explicit irrevocable answers constitute what can be tested? Even scientific models, such as electron transport systems in photosynthesis and cellular respiration, are presented as facts. Yet what is intended to be a factual representation in one text is often interpreted differently or disputed in another text. But rarely are students given an opportunity to examine alternative explanations. In most classrooms textbooks are expected to be the harbingers of factual knowledge. What is written should be explicit, precise, and limit student interpretations. Alternative frameworks are often regarded as something to be corrected and equated with misconceptions.

Like textbook presentations, standardized testing attempts to negate ambiguity. However, the process of scientific investigation suggests that ambiguity is a constant companion (Bronowski, 1981; Kuhn, 1970). In critiquing rules for uncovering the truth held within scientific explanations Polanyi explains:

> Indeed, when we apply any of these formulations for deciding a great question in science, we find that they prove ambiguous precisely to the extent of allowing both alternatives to be equally arguable. (Polanyi, 1962, p. 165)

By imposing an objectivist approach that looks closely at the completed task rather than the subject, we confine what is assessed. The student with 70% soon becomes the 70% student. In the most extreme situations, the 70% may even stand in place of the student. Their abilities, interests, and attitudes may even be defined by a number.

The idea that 50% students have difficulty organizing their portfolios was proposed at one of the action research meetings. When one teacher suggested that some students, who achieve an honors mark on exams, might also have some difficulties organizing their work, the query was not addressed. After the meeting I asked the teacher, who proposed the concern about what he referred to as 50% students being disorganized, whether if it was possible that all 50% students would have difficulties assembling their work or that all honors students would have few difficulties. His quick answer was that only rarely would this happen. Predictability was used to legitimize his personalknowledge.

Saul (1992) argues that a truth built upon certainty requires the virtue of repeatability to ensure predictability. However, he argues, this preoccupation with certainty moves the focus of knowledge toward events rather than people. Sunrises are predictable; however, the meaning you construct or the emotion you feel while viewing the sunrise is not repeatable. Not every sunrise means the same thing to us throughout our lives. Similarly, learning poetry or viewing art is not understood by repeatable performances. Eisner (1985) embraces ambiguity because it shifts the focus from the products of human interaction to the process.

> I welcome ambiguity and uncertainty because I believe that, the quality of inquiry is as least as important as arriving at the church on time. (Eisner, 1985, P. 2)

The humanist alternative

An alternative to the rationalist paradigm has been identified as the humanist orientation (Langer, 1957; Polanyi, 1962; Ashton-Warner, 1963; Dennett, 1991; Saul, 1992; Gadamer, 1993; Dimasio, 1994b). The humanist orientation focuses upon the subject and object as part of the interaction. The purpose of the humanist approach is not to reduce ambiguity and distill a pure, single answer, but to report the complexities of interaction and interpret what is being said in each case. Gadamer explains that:

> ... an interpreter's task is not simply to repeat what one of the partners says in the discussion he is translating, but to express what is said in the way that seems most appropriate to him considering the real situation of the dialogue,

> which only he knows, since he alone knows both languages being used in the discussion. (Gadamer, 1993, p. 308

Eisner (1985) compares the scientific, technological approach with the humanist approach which he explains arises from an artistic orientation.

> Scientific activity yields propositions so that truth can be determined in relation to its instrumental value, a value dependent upon its predictive or explanatory accuracy. Artistic activity creates symbolic forms which themselves present directly an idea, image, or feeling which resides within rather than outside of the symbol. (Eisner, 1985, p. 89)

The purpose of the humanist approach is not to supplant or replace the scientific approach but to provide another way of viewing and interpreting social life. Eisner explains:

> Any framework, any representation, and any methodology has limited parameters. Once it is granted that modes of knowing and the forms through what one knows is represented are multiple, it makes no sense to restrict inquiry to a single form. (Eisner, 1985, p. 7)

The humanist approach to understanding assessment is not designed to reduce the complexities of student understanding to a single indisputable answer, or interpret student progress by a number or letter score. Eisner insists that:

> Rather than reduce the human mind to a single score, qualitative inquirers attempt to adumbrate its complexities, its potential, and its idiosyncrasies. (Eisner, 1985, p. 140)

Eisner explains that the dominant, reductionist approach to educational inquiry, the scientific approach, is supported by quantitative data. The focus has been placed on outcomes that can easily be monitored rather than the experiences of the learner. What is most needed is not the abandonment of the scientific approach but another perspective from which to view assessment. We must know what students can do, but as educators we should be equally interested in learning how these tasks were accomplished.

> Schoolmen have been woefully derelict in giving the public anything other than standardized methods for appraising educational quality. There have been few alternatives to highly reductionist indices of learning available. By developing, not so much alternatives, but complements to conventional approaches to evaluation and research, the possibility of balance in view, in method and in 'data' can be created - at least in principle. ((Eisner, 1985, p. 139)

Qualitative inquiry provides a sharp contrast to the dominant scientific approach. The qualitative inquirer investigates meaning rather than behavior. A single act could be performed for many different reasons.

> To the qualitative inquirer ... one must try to uncover the meaning of action, moves, behaviors and not simply the fact that behavior has occurred.
> The qualitative inquirer is likely to be interested in the meaning of the move perhaps even more than the move itself. Thus he is in a position to explicate the costs as well as the benefits of certain forms of achievement. By attending to meaning rather than behavior as such, by relating behavior to culture, and schools, the qualitative inquirer is in a position to secure a much more complex view of educational situations. (Eisner, 1985, p. 140)

In sharp contrast to the scientific approach, validity is not measured by repeatability. The acceptance or rejection of an interpretation is not measured by predictability. The experiences of one student are not used to represent all students. The strength of the value of the interpretation is evaluated by the manner in which it allows us to understand the story. Gadamer explains:

> The individual case does not serve only to confirm a law from which each practical prediction can be made. Its ideal is to understand the phenomena itself in its unique and historical correctness. However much experimental universals are involved, the aim is not to confirm how men, people, and states evolve – but to understand how this man, this people, or this state is what it has become, or more generally, how it has happened that it so is. (Gadamer, 1993, p. 5)

Eisner (1985) writes of the interpretations of learning events as educational criticism. Not criticism in the commonly used sense, arising from diagnosis. But criticism that arises from telling and bringing about change.

> ... good educational criticism, like good criticism of anything else, should help the reader or listener see more than he or she would without the benefit of the criticism. (Eisner, 1985, p. 101).

The movement toward the descriptive interpretation of events provides another way of understanding assessment.

> I would argue that we support and expand the current effort being made to broaden the ways in which we evaluate. Conventional modes of evaluation, particularly the use of achievement tests, are designed to capture only a slender slice of educational life. ...
>
> To use such tests as exclusive tools is like casting a net into the sea that is intentionally designed to let the most interesting fish get away. (Eisner, 1985, p. 176)

Notion of connoisseurship

In order to bring an understanding to the complex world of assessment, I began by looking at the notions of connoisseurship as explained by Eisner (1985) and Polanyi (1962). Most of the decisions that teachers make are based on connoisseurship. Mehrens (1992) explains that for new forms of assessment, such as performance assessment, there is heavy reliance on observation and professional judgment. Personal knowledge, gained from interpreting the experiences of students as they learn and complete assignments, provides the greatest knowledge base. Little time is available in day-to-day teaching to take formal surveys to determine student opinions or interview students about their reactions to the work. Information is most often gathered within the context of the lesson. Teachers within the action research group did not talk about the data derived from diploma exams as a means of understanding their assessment techniques or the students in their classes. Rather, they spoke of the individuals in their classes as a means of making sense of the data. It is here that dichotomy between teachers and many school administrators became visible. Administrators tended to look at exam results to identify "good teachers" and "good students". Teacher practical knowledge focused on the subjects as a way of gaining knowledge, while administrators focused on objective knowledge supplied by exam results.

Eisner (1985) explains connoisseurship in education, as in other areas, as the art of perception that makes the appreciation of such complexity possible. Polanyi (1962) describes connoisseurship, as a skill, that can be communicated only by example and by precept. Polanyi forges a link between personal knowledge and connoisseurship by explaining:

> Personal knowledge is manifested in the appreciation of probability and of order in the exact sciences, and see it work even more extensively in the way the descriptive sciences rely on skill and connoisseurship. At all these points the act of knowing includes appraisal; and personal conjunction between subjectivity and objectivity. It implies the claim that man can transcend his own subjectivity by striving passionately to fulfill his personal obligations to universal standards. (Polanyi, 1962, p. 17)

Criteria for authenticity

Wiggins provides criteria for authentic testing. He explains that "evaluation is typically based on judgment that involves multiple criteria (and sometimes multiple judges), and that judging is made reliable by agreed-upon standards and prior training" (Wiggins, 1989, p. 711). Unlike traditional assessment, arbitrary boundaries, such as time constraints, do not limit assessment. Wiggins (1989) indicates that portfolios or a full season's schedule of games provide a more accurate picture of student progress. To be authentic the assessment practice must not be structured to provide summative evaluation but to provide the needed information to direct learners toward a mastery level.

To be more authentic, alternative assessment must do a better job of informing the learner about the manner in which he or she learns. Wiggins (1989) warns that tests which continue to perpetuate the blind use of algorithms do little to promote students' true understanding. In addition, "authentic tests are contextualized, complex intellectual challenges, not fragments and static bits or tasks. They culminate in the student's own research or product, for which 'content' is to be mastered as a *means*, not as an *ends*" (Wiggins, 1989, p. 711). Unlike traditional tests that utilize aggregate grades and restrict the number of variables, authentic tests use multifaceted scoring systems that attempt to take into account multiple variables that reflect learning. The teacher's role also changes in authentic assessment. Rather than suspending personal judgment to objectively tally grades, the teacher relies heavily on his or her judgment, as a connoisseur might, to assess the many, complex indicators of achievement. The judgment of the teacher is called into play. The students, because they become intimately involved in the selection of the material to be evaluated, employ judgment and become active participants in their own assessment. Wiggins (1989) explains that the attempt to examine the entire context of learning in all of its complexity, as opposed to isolated components, reduces certainty. He explains:

> They (authentic tests) are designed to emphasize realistic, but fair, complexity; they stress depth more than breadth. In doing so, they must necessarily involve somewhat ambiguous tasks or problems, and so they make student judgment central in posing, clarifying and tackling problems (Wiggins, 1989, p. 711).

The movement toward the assessment of knowledge that resides within the subject, as opposed to objective knowledge, changes the focus the criteria for grading and the reporting of grades. "Authentic tests measure essentials, not easily counted (but relatively unimportant) errors" (Wiggins, 1989, p. 711). What the student can accomplish becomes of primary focus, not grading curves or the analyses that attempt to compare performances of different students for the purpose of selection or elimination. Wiggins explains:

> Rather than rely on right/wrong answers, unfair 'distracters' and other statistical artifices to widen the spread of scores, authentic tests ferret out and identify (perhaps hidden) strengths. The aim is to show students what they can do. (Wiggins, 1989, p. 711)

More authentic assessment must allow for different types of learners by providing choices of assessment (Herman, 1992b; Hofstein, 1988; Joint, 1992; Klainin, 1988; Maeroff, 1991; O'Neil, 1992; Stake, 1972). Time required, location and support resources must be addressed. Wiggins raises the following questions:

> Why must all students be tested in the same way and at the same time? Why should speed of recall be so well-rewarded and slow answering be so heavily penalized in conventional tests? (Wiggins, 1989, p. 712)

Summary table of traditional and alternative assessment strategies

Traditional view and expectations	Emerging view and expectations
• Reliability primary goal	• Validity primary goal
• Knowledge and application of knowledge is the primary focus.	• Interpretive and communicative skills stressed. Creative and critical thinking skills, personal values and social skills component acknowledged.
• Subject matter is defined by what can be tested. Unambiguous answers are sought.	• Ambiguity and doubt are acknowledged in finding answers. Limitations of conclusions are acknowledged. Certainty is not the primary goal.
• Information is decontextualized and presented objectively. Isolated questions are the norm.	• Knowledge is contextualized and linked with attitudes of the learner. Assessment is often done by way of case study, journal entry, by peer assessment.
• Multiple choice format predominates and testing is based on the principles of standardized assessment.	• A wider scope of assessment strategies, that attempt to link the context of learning with assessment. Group assessment and self assessment become part of the emerging strategy.
• Assessment founded upon the notion that every student needs to be assessed on the same material.	• Assessment strategy acknowledges that assessment, like learning, is idiosyncratic.

Finding a direction for change

Embedded within the question, "what is wrong with current assessment practices?", is the assumption that some commonalty of what good assessment should be exists. Surveys of assessment practices (Doran, et al., 1994; Douglas, 1983; Hofstein, 1988; Klainin, 1988) reveal that although certain components are easily identifiable, there is no true consensus of opinion about how students should be assessed or what should be assessed. What is currently done, according to many researchers, is motivated primarily by economics (Popham, 1993) and intended more for selection than educating (Mehrens, 1992).

The myth that we can progress begins with the assumption that everyone is pointed in the same direction and is using the same compass. What makes the diversity of opinions about what should be evaluated (valued) even more incredible is that there would appear to be tremendous consensus about what is most important in science.

A number of curriculum guides (Alberta, 1991; New Brunswick, 1992; Newfoundland, 1990; Nova Scotia, 1993; Ontario, 1987; Quebec, 1987) provide goal statements that closely parallel the reasons to study science, as identified by the Science Council of Canada report, *Science For Every Student: Educating Canadians for Tomorrow's World* (Science Council of Canada, 1984). Each of the curriculum guides acknowledges that students need a science education that will "(1) develop citizens able to participate fully in the political and social choices facing a technological society; (2) train those with a special interest in science and technology for further study; (3) provide an appropriate preparation for the modern world of work; and (4) stimulate intellectual and moral growth to help students develop into rational, autonomous individuals" (Science Council of Canada, 1984, p. 13). Yet, in spite of the overwhelming consensus about why students should study science, each of the curriculum documents concentrates most heavily on the presentation of content. None of the curriculum guides provides a link between the intent of the curriculum and the manner in which student are assessed.

If the goals stated by the former Science Council of Canada (1984) are valued, how are they reflected in the manner in which students are assessed? Has evaluation (or what we value) really changed amidst a plethora of curriculum revisions or does acquisition of scientific knowledge for further study, as assessed by multiple choice exams, continue to overshadow the other reasons to learn science? Do our assessment practices demonstrate that we value scientifically literate citizens, that we have a commitment to prepare students for the modern day work world, or that we foster the intellectual and moral growth of our students?

Obstacles to change

Progress toward developing alternative forms of assessment has been fueled by a desire to provide a better match between student learning, curriculum presentation, and assessment (Collins, 1992; Darling-Hammond, 1991; Eisner, 1993; Maeroff, 1991; Shavelson, 1992). The sometimes incongruent aspects of curriculum arise because they are supported by different epistemological frameworks, research traditions, and orientations toward teaching and learning (Ryan, 1994; Simmons, 1992; Tierney, 1991; Wiggins, 1992a). The movement in assessment that has attempted to harmonize teaching and assessment arises from critical reflection about current classroom practices and originates from teachers' practical knowledge. The experiences are context specific and the knowledge derived is inextricably tied with the subject matter, the children involved, teaching strategies employed, and perceptions of the school community.

A second but distinct movement in assessment, driven by a political agenda, demands accountability. All too often this agenda is viewed as an ally because it identifies many of the same problems with traditional student assessment. However, the accountability movement is founded upon a hierarchical organizational model that is motivated by surveillance, comparison of results, and remedying deficient situations. Subsumed within the accountability movement is a corporate agenda more interested in measuring student achievement to identify competency, than questioning the manner in which students are being assessed (Eisner, 1985). The assessment methodology is not deemed as problematic. No one asks if these tests provide a proper indication of what students learn and the manner in which students learn (Wiggins, 1989). By accepting the taken-for-granted assumption that the tests are valid, student achievement is used to compare and rank teachers, and schools. Good schools and good teachers are the ones whose student performance is consistently above the mean. Accountability becomes a way of measuring economic input (the money invested in education) with what economists refer to as output (what learning has been accomplished). Saul quotes Geraldine Gilliss:

> Slow growth, unemployment, and fierce international competition have led to a resurgence of neo-conservative philosophies concerning the dominant role that private institutions should play. There is a general trend toward catering to the interests of private business and industry, which are viewed as both the source of future prosperity and the sectors most likely to provide needed employment. (Saul, 1992, p. 5)

The *International Comparison in Education : Curriculum, Values and Lessons*, (1992), produced jointly by the Alberta Chamber of Resources in partnership with Alberta Education exemplifies how the corporate agenda has raised concerns of accountability above those of matching learning and assessment. Great effort is taken by the authors of this document to demonstrate that Alberta's students do not do well on international mathematics and science exams without ever making the type of exams problematic. By tying what the authors perceive to be inferior mathematics and science education to a downturn in the economy, they draw upon the growing paranoia and conclude that education is ultimately responsible for the economy. A comparison of the conceptual level of science and mathematics education is provided (1992, p. 23-27) to underscore why Canada is losing its economic place in world trade. Conveniently, the authors fail to explain why some of the countries cited for conceptually advanced mathematics and science, such as Czechoslovakia and Hungary are experiencing economic difficulties that far exceed those of Alberta. Could it be that education is not a single factor that is responsible for a healthy economy?

Administrators, teachers, parents, the community, and children ask different questions about assessment. Unfortunately, people often seek to answer all of the questions with one type of information (Smith, 1991). The information provided by the rationalist paradigm not only defines what questions can be asked, but also ensures that answers come from the same domain of understanding.

The reasonably unified displeasure with current assessment practices should not be confused with a concordant vision of where we should begin moving. Nor can we begin to view alternatives until we begin thinking about students, teaching and learning within a new light that focuses upon the individual subject and the context in which learning takes place. The humanist alternative is not founded upon constructing a method, transferable to everyone within a classroom, but in developing a sensitivity for interpreting what has been accomplished and analyzing what the individual still must do in order to achieve an acceptable standard of performance.

The movement toward alternative assessment: England

The movement toward alternative forms of assessment has grown fastest and has had the greatest time to mature in England. In part, the movement toward authentic assessment was spurred by dissatisfaction with a long-established testing tradition (Baird, 1986; Berlak, 1992; Darling-Hammond, 1991; Dorr-Breme, 1983; Eisner, 1985; Herman & Golan, 1992; Jackson, 1968; O'Neil, 1992; Stake, 1972). The focus of what came to be known as "authentic" assessment was placed upon complex performances, often described as the three P's - performance, portfolios, and products. The use of portfolios in English, some elements of investigational work in mathematics, oral work in modern languages, and extended projects that involve research, reflection, and writing in the humanities and sciences signal attempts to harmonize learner tasks and assessment practices. Nuttal (1992) indicated that, at the grade 11 level, most of the work is graded by the candidate's teacher, but is subject to a monitoring board. Monitoring occurs throughout the year by a part-time moderator who visits schools.

In 1988, the British Education Reform Act introduced a national curriculum. Prior to the Reform Act, primary schools had tremendous autonomy and were charged with determining their own curriculum. Secondary schools had their curricula defined largely by independent examination boards. The examination boards legitimized student learning by identifying students who had mastered specific content and skills at the end of their high school tenure. The focus of the examination boards was not on developing curricula but the selection of students for post-secondary studies (Madaus & Kellaghan, 1989). In 1988 the assessment of students, under the guise of the Educational Reform Act, established attainment standards and extended assessment to students in grade 2, 6, 9 and 11 GCSE (Graduate Certificate for Secondary Education).

An ad hoc committee was struck to establish "standard assessment tasks" or SATs. According to Nuttal (1992) the standard attainment tasks provide teachers with a new lens for viewing student learning and help teachers understand how different students learn. The new assessment tasks were also designed to provide a positive environment for viewing student work that integrates testing with everyday classroom experiences (Madaus & Kellaghan, 1989). The task group only provided general recommendations and allowed other bodies to define the framework for assessment. However, the task groups did recommend the exploitation of varied forms of presentation, such as oral, written, pictorial, video, computer, and practical demonstration(Madaus & Kellaghan, 1989). In 1991, in England and Wales, national testing was provided for grade 2 classes in mathematics and science.

Although the focus of the testing was to provide diagnostic information and formative feedback, the tests were often used to compare schools to promote free market competition for students. The notion that schools, curriculum presentation, and teacher effectiveness could be assessed by wider-scope exams emerged. Madaus and Kellaghan (1989) quote from the British Education Report on Assessment (BERA) ". . . the reality has been an emerging system with a more clearly defined emphasis on the use of assessment for the comparison between schools and the generation of a national picture of education." (1989, p. 462) By sharp contrast, in Scotland, the exam results were not published and only used for the diagnosis of student problems. Scottish teachers viewed the exams as having low stakes, while those in England and Wales were viewed as high stakes.

Several management problems were identified. Assessment procedures are extremely time consuming. Nuttal (1992) reports that an average of 44 hours was required. Madaus and Kellighan (1989) indicate that for some tasks more than 6 weeks of class time had to be allocated for testing. In addition between 82 to 90 hours to plan for the assessment, collect needed materials for administration, mark the tasks, and record the marks is required (Madaus & Kellaghan, 1989). In most cases another teacher is required to work with the class as small group testing occurs. Madaus and Kellighan explain:

> Many teachers reported that the SATs were almost impossible to administer unless another teacher assisted or part of the class was taken by an auxiliary teacher. Thus additional staff members, parents, volunteers, and other resources had to be found in many schools. (Madaus & Kellaghan, 1989, p. 462)

The procedure is also extremely expensive. One estimate pegged the development costs to run at more than £6 million, with another £60 million for administration (Madaus & Kellaghan, 1989). Popham (1993) provides a detailed accounting of the costs associated with conducting large-scale authentic testing. Teachers questioned the value of testing and many stated that the monies would be better served on improving instruction rather than monitoring students and teachers (Klainin, 1988; Madaus & Kellaghan, 1989; Nuttal, 1992).

Teachers also expressed concerns about the daily activities of children who were being assessed (Madaus & Kellaghan, 1989). The loss of daily contact greatly disrupted normal classroom practices and caused discipline to deteriorate in younger grades. In addition the disruptions caused by testing caused the rescheduling of meetings, reallocating of space and resources, and extra planning for teachers.

Madaus and Kellighan (1989) reported increased teacher stress as a negative byproduct of the foray into national testing. Not surprisingly, teacher stress was greatest in England and Wales, where testing was thought to provide greater accountability. Aside from the organizational problems created by the demands of increased testing, teachers also dealt with the anxiety of parents who worried about how their children were adapting. Yet other parents used testing scores to check up on teachers "and as a mechanism of getting rid of the bad ones." (Madaus & Kellaghan, 1989, p. 463).

Some resistance was also noted by teachers who believed that the testing situation was not rigorous enough. According to Nuttal (1992), this minority view had the ear of John Major and gained considerable support from lobby groups concerned about the competition of England's students in the European Marketplace. In 1992, the Standard Assessment Tasks (SATs) were canceled for grade 9 and significantly reduced for grade 2, so as to be accomplished by one teacher. The number of tasks has also been reduced to take less time. An economic agenda, not pedagogy, dictated long-lasting change.

Perhaps the most important information gained from the British experience is derived by comparing the experiences in Scotland with those in England and Wales. In England and Wales, where "high stakes" assessment was employed, SATs were seen to foster consumerism, rather than promote educational discourse (Madaus & Kellaghan, 1989). The adoption of large-scale performance-based-assessment strategies is not immune to many of the problems surrounding the employment of large scale multiple choice exams (Mehrens, 1992). Madaus and Kellaghan (Madaus & Kellaghan, 1989) explain:

> Further, using the SATs in a high stakes context was viewed as more likely to have an adverse effect on both teaching and learning, by narrowing teaching to assessment. (Madaus & Kellaghan, 1989, p. 468)

In Scotland, where the focus of assessment had been placed on manageability and validity, the primary concerns were placed on helping students achieve the indicators. In England and Wales little time was spent defining manageability and validity, these tasks were left to the teachers. Robust data for was generated for the purposes of comparability. Validity lost out to reliability (Madaus & Kellaghan, 1989). Similar negative results of using alternative assessment for large scale testing have been recorded in the United States (Darling-Hammond, 1991).

> Tests that are externally designed and imposed can never play an important role in school improvement, since they deny teachers and students the opportunity to be a part of the process of developing and wrestling with standards. (Darling-Hammond, 1991, 224)

The movement toward alternative assessment: Canada

A wide variety of alternative assessment projects have been initiated in Canada during the 1990s. Like England, the most visible projects are those instituted at a National or Provincial Level. Because most of the large-scale projects are designed to produce reports rather than provide teachers with alternative strategies, the parameters of a standardized assessment model are strictly followed.

The Science Assessment Project (Council of Ministers, 1994, 1995, 1996) was instituted to provide a profile of student achievement across Canada. The Science Assessment project is only one component of the larger School Achievement Indicators Program or SAIP, in which student achievement is monitored in mathematics, reading, writing, and science. Providing individual student assessment is not a goal of the project. Anonymity of students and school jurisdictions, according to the Information Bulletin (Council of Ministers, 1996), will be preserved. Reports will present indicators of the knowledge, skills and attitudes that students possess in various parts of Canada. Student achievement will be grouped into five categories according to mastery of the skill or concept and results will be reported in terms of the percentage of students attaining each of the five performance levels. Sixteen and thirteen year olds were chosen for SAIP. Although comparisons of different provinces and the science curriculum in the provinces was not set as a goal, concerns were voiced at an ATA Science Conference in October, 1995 that such an agenda had infiltrated the project. Performance-based assessment is outlined in the Information Bulletin (1995) as a methodology for assessing a wider range of science skills.

The Science Assessment Project of the School Assessment Indicators Project departs from traditional assessment in that students are required to demonstrate an acquisition of science inquiry skills as they perform tasks. According to the information Bulletin (1996):

> Students participating in the science inquiry skills assessment will perform various tasks that require them to generate and analyze their own data by applying scientific inquiry skills to answer questions of a scientific, technological, and/or societal nature. (Council of Ministers, 1996, p. 5)

The two-hour test sets a context for student testing by indicating that they are a researcher for a toy company that is developing a game involving a bouncing ball (Council of Ministers, 1995, p. 14). The student is asked to test various balls and compare their ability to bounce. Data collection, hypothesis formation, the control of variables, identification of error, and general principles of experimental design are assessed for individual students.

Alberta Education (1993, 1994, 1995) has randomly monitored the development of science skills, by way of performance-based assessment for grade 11, 9, 6 and 3 since 1993. The laboratory-directed approach to assessment is referred to by the acronym CAMP (Curriculum Assessment Materials Project). Inquiry tasks are provided and evaluated holistically on a four point scale: 0 = misunderstood the task, 1 = not yet at grade level, 2 = at grade level, and 3 = above grade level. Booklets of sample responses have been provided to give teachers measurement indicators for each of the defined levels. Unlike the SAIP, Science Project, the provincial approach of CAMP provides a model for teachers to emulate. Up to 1995 teacher involvement has been restricted to a volunteer basis. Teachers wishing field tests have the option of signing up for them. No formal reporting of individual, class, of jurisdiction marks have occurred to date.

The primary goal of the CAMP field test is not in developing a model to assist teacher assessment, but in refining a process that will permit random monitoring of science skill development. Teacher involvement and input in the project has been compromised because they recognize that this model acts as a prelude for compulsory exams in the future. The impetus for modifying the CAMP model to better reflect idiosyncrasies of an individual classroom is slowed when teachers recognize that it is being field tested to provide a universal model, suited for all classrooms.

The British Columbia Assessment Report of 1986 (Ministry of Education of British Columbia, 1991, Ministry of Education of British Columbia, 1986) set a new direction for broadening the parameters for student assessment. Science concepts, the nature of science, science and technology, science-related careers, and science and society were all considered as important components of science curriculum and worthy of assessment. The authors of the Technical Report II: Student Performance Component state:

> As a project team we view assessment of performance as one way to assess the practical intelligence of students. Like many others we were and continue to be dissatisfied with testing which calls on a very small range of student abilities to represent knowledge in symbolic form. (Ministry of Education, Province of British Columbia, 1991, p. 254)

Unlike the CAMP materials developed in Alberta, the Assessment Project in British Columbia concentrated on developing sample activities with accompanying assessment strategies that would enable teachers to expand their repertoire of assessment strategies to better represent student learning in all aspects of science. In order to survey as many content topics as possible, three parallel test forms were developed at each grade level. Each test consisted of affective and achievement test items linked to a curriculum objective. Each student worked independently to complete only one of these individual tests. The affective component of the survey included items related to school science, science careers, science in society, and special issues. The major domains addressed with the achievement tests were science processes, knowledge, higher level thinking, technology, and the nature of science.

In 1986 and again in 1991 the student performance component provided 13 stations of tasks organized at the grade 4, grade 7 and grade 10 levels (Ministry of Education, Province of British Columbia, 1991). Tasks at each of the stations were designed to take between 6 to 8 minutes . Students moved from station to station as individuals. The results of the 1991 performance testing in British Columbia were included in a separate report (Erickson, et. al., 1993).

Rather than being decontextualized, the skills were designed to reflect core science curriculum and the skills that were being used at the time. The primary focus of the tasks became a way of interpreting curriculum rather than monitoring teaching or student achievement. The British Columbia performance assessment was based on the following six Science Skill Learning Dimensions:

1. Observation and classification of experience
2. Measurement
3. Use of apparatus
4. Communicating
5. Planning Experiments
6. Performing Experiments

The Ontario Ministry of Education (1981, 1983, 1989) sponsored the development of various items for all high school science courses, entitled Ontario Assessment Instrument Pools. The first document was physics, developed in 1981, chemistry in 1983 and finally biology in 1989. The exam bank was developed as a way of assessing learning outcomes specified in curriculum documents. Because standardized assessment was not the goal, exam blueprints and reliability standards were not established. The initial focus of the exam items was in providing teachers with support for determining whether students had achieved learner outcomes specified by curriculum documents. Although the majority of items were of the pencil and paper variety, some evidence of laboratory-related questions and laboratory tasks can be found in the various volumes. Nowhere in Canada is the approach to assessment less centralized. A number of school districts have assumed leadership for the development of consortiums that have attempted to expand assessment strategies (Midwood, O'Connor, & Simpson, 1993).

Summary

The standardized assessment model gains acceptance because it is supported by a framework of rational thinking. An assessment model formulated on rationalism sets reliability as a primary goal. Repeatability and the ability to make predictions provide value. Standardized assessment is founded on the philosophical orientation of rationalism. The humanist alternative recognizes validity as a primary purpose. Here validity is achieved by linking the context in which students are assessed to the idiosyncratic manner in which they learn. Not every student is treated in the same manner or assessed in the same way. Assessment strategies that attempt to contextualize learning are considered "authentic".

Historically, thinking considered non-rational or irrational has been given lower status, yet many decisions are made on the basis of emotion, intuition, or personal knowledge. Assessment strategies founded upon rationalism are designed to eliminate uncertainty. Experts are defined by the exactness of answers. A 70% is decontextualized from classroom, student and subject. Many assume the grade indicates what percentage of the curriculum that students understand, rather than what mark they achieved on the test. Authentic assessment, by contrast, recognizes the ambiguity and doubt as a part of the journey.

Chapter 3: Adopting a research approach

Overview

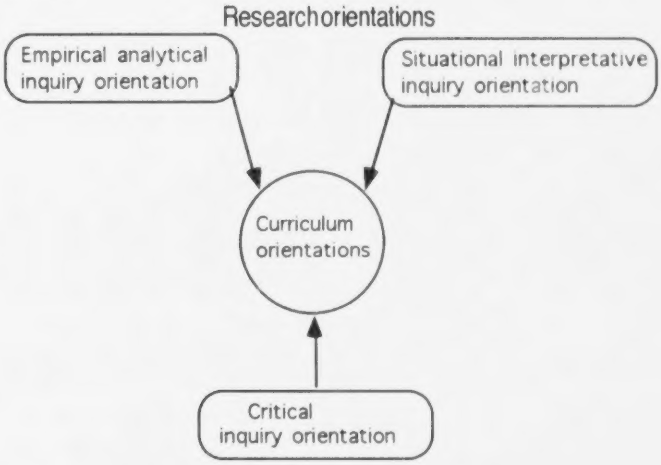
In this chapter I will explore three orientations to research as described by Aoki (1978). The potentials for developing an understanding of the research questions and limitations inherent within each orientation are examined. Within the chapter I explore the reasons for choosing an action research orientation. A definition of action research is provided along with its limitations.

In the second part of the chapter, I explore why it is important to bring about change from within the school, as opposed to imposing change upon schools. Finally, I describe the origins of the action research group and how I came to understand the research questions. Schwab's tables of inventions (Kemmis, 1990) are used as a means of organizing research questions. The table of inventions is based on the four commonplaces of education: teachers, students, subject matter, and milieu.

Research Orientations

The research questions posed can be categorized into three inquiry orientations described by Aoki (1978) as empirical-analytical, situational interpretative, and critical theoretical.

Figure #1: Research questions


Research orientations
Empirical analytical inquiry orientation
Situational interpretative inquiry orientation
Curriculum orientations
Critical inquiry orientation

The empirical orientation is a technical approach that addresses problems defined by " ends-means". Within the context of assessment this orientation concentrates on providing an efficient methodology or means for achieving curricular/evaluation goals and objectives. Questions which address congruency between curriculum intent and assessment strategies would come under the ends-means approach (Aoki, 1984). The question of curriculum and evaluation fit is a primary concern raised by the teachers involved in the seminars and meetings; however, other questions within the research extend beyond the technical, ends-means orientation. The technical orientation accepts the scientific cannon of repeatability and reproduction. Quantitative methodologies, such as survey techniques, are often applied to these questions and data are most often collected and interpreted in terms of reliability. The technical (ends-means) approach to evaluation is generally well accepted because it is congruous with the manner in which curriculum is developed. Aoki emphasizes these two factors when he states:

> These end-means concerns reflect an orientation to evaluation which can be characterized as technical or instrumental. As such these concerns reflect the dominant evaluation approach in use, going hand-in-hand with the technically oriented mainstream curriculum development/evaluation rationale, known popularly as the Tyler Rationale. (Aoki, 1984, p. 7)

Survey methodology has been a long-established technique to determine how curriculum changes were reconciled by teachers. One such survey, conducted by Douglass and Kahle (1983), explored the amount of time biology teachers believed that they spent on lectures, discussions, laboratory activities, and field trips. Teachers' perceptions were compared with students' perceptions and the evaluation outlines used assess students in each of those areas. Teachers' perceptions, students' perceptions, and observations by researchers did not match. Most importantly, examinations did not reflect the time spent in various activities. Lectures provided the greatest amount of testable information. Assessment studies, such as Stake (1972), support the conclusion that information-based knowledge is most often assessed. Douglass and Kahle (1983) found that the respondents relied most heavily on what was defined as "objective exams" in determining a student's grade. Unfortunately, the researchers made no attempt at defining what they meant by "objective exams" or how the teachers may have interpreted their meaning. The researchers concluded that the teachers valued tests as a primary evaluation tool. These results are supported by other findings (Anderson, 1989; Dorr-Breme, 1983; McLean, 1985; Wilson, 1989) which show a preoccupation with formalized testing by paper and pencil exams as a practical evaluation tool.

For my study a student survey was constructed to provide an overview of student reactions to changes that were attempted in assessment practices. A list of questions was constructed from a series of initial interviews with teachers and students. The survey was only used to provide the action research group at Eastern High School with quick feedback during the developmental phase. Did students agree with teacher perspectives of teachers, or, as Douglass and Kalhe (1983) found, would student responses indicate important differences in interpretation? Trends and statements obtained from specific students could be surveyed for commonalty of acceptance and act as a springboard for the creation of other interview questions.

Although the survey methodology provides a panorama of many opinions, the survey fails to answer many of the deeper questions. For example, do students believe that factual knowledge is the most valued? Do students and teachers share common meanings for terms such as: "factual knowledge", "objective" testing, and "science-technology-society". A most important question, "Do teachers or students have a concern about what is *valued* in evaluation strategies?", can not be addressed by surveys. Survey methodology is apt to show trends based on cause-and-effect relationships, but provides little opportunity to question the tenets of the cause-and-effect assumptions.

Critical theorists have expressed concerns about ends-means approaches that examine the products of human interaction, while neglecting the processes by which these interactions take place. Elliot points out that "the emergence of heirarchialized, specialist functions to control and regulate primary practice is characteristic of centralized and technocratic systems of schooling" (Elliot, 1991, p. 55). He argues that this culture largely supports a non-reflective, intuitive, and highly routinized form of practice. The technocratic system, supported by the technical research orientation (ends-means orientation), is directed toward hierarchical surveillance and control over teachers. Outcomes or products are defined without any true investigation of the processes of change. By subordinating process to product, any knowledge of process is derived largely from inference. The ends-means orientation seeks to eliminate variables to identify single cause-and-effect relationships. Although the notion of one cause and one effect has value in the scientific world, it is severely limiting when studying human interaction.

The situational-interpretative orientation has been used to formulate meaning. A variety of research methodologies can be used to identify how teachers use practical knowledge to choose certain evaluation strategies. Duffee and Aikenhead (1992), used an ethnographic methodology, to explored how teachers make decisions concerning student evaluation of new content associated with teaching science through a science-technology-society (STS) orientation. The researchers applied the heuristic model of teacher practical knowledge (Lantz & Kass, 1987) which enabled them to reflect upon how teachers make decisions about evaluation practices.

Duffee and Aikenhead (1992) drew support for their initial assumption, that teachers adapted curriculum in ways they think are most appropriate for specific teaching situations, from previous research. These adaptations were derived from feelings and intuitions gathered from their life experiences as both student and teacher. In addition, the researchers drew upon other work to explain that these adaptations were influenced by content of the curriculum and the context of the teaching situation (Jackson, 1968; Lantz & Kass, 1987). The work of Feimer-Nemser and Flood (1986) suggests that the decisions for curriculum implementation, within the classroom, are situationally specific, personally compelling, and oriented toward action. Teachers use practical knowledge to gain insight into practical knowledge that informs the decision-making process.

Duffee and Aikenhead (1992) choose student evaluation strategies as a means of studying how teachers use practical knowledge to make decisions about new curriculum. The research is founded upon a second assumption; that teachers give high priority to the issue of student evaluation when they contemplate the adaptations of a new curriculum (Aikenhead, 1984).

Hofstein (1988) used a similar methodology to conduct a study in which teacher practices for assessing the laboratory component of senior science courses were investigated. Observers were placed in classrooms to record teacher practices. Methods of assessment were categorized as post laboratory write-up, practical examination of task, and continuous assessment. Observers were to classify assessment practices according to pre-established groups and record, by way of interview, teacher reasons for the assessment strategy chosen. Hofstein reported that a narrow band of assessment techniques dominated observed classrooms.

Although the interpretive approach provides a record of events as teachers cope with changes, the approach does little to assist teachers who wish to change their practices. The observer is detached from the work of the teacher. Providing a record of divergent practices may enable greater understanding, but it does little to provide counsel for social action, a primary concern for the teachers in the group that I began working with. In ethnographic studies lived-experiences are described, interpreted, contextualized, and sometimes evaluated, but rarely do interviewees find problem-solving strategies (Miles & Huberman, 1984). The agenda is set by the researcher and the process of analysis is often conducted by the researcher in isolation. The ethnographic orientation makes no claims of social change (Aoki, 1978).

Interpretive studies are constantly in danger of arrogating teacher knowledge, without providing teachers with any direct benefits. The moral decision to restrict the study to an interpretative orientation would appear clear. The interests of the teachers did not reside with describing the discontinuity between curriculum and evaluation, but in changing their evaluation strategies. Prior to beginning my study, teachers acknowledged a general dissatisfaction with their assessment practices.

The third orientation to curriculum inquiry, described by Aoki (1978), is the critical orientation. A primary interest of this orientation is improving the human condition by rendering transparent tacit assumptions and by initiating a process of emancipation. Knowing, in this orientation, can be described as critical knowing that is derived from reflection and action. Understanding is derived from what Schon (1983) calls reflective practice. Kemmis and McTaggart point out the tensions that occur when educators begin to reflect on pedagogy from a critical perspective:

> A critical and self-critical approach to education like that of other forms of educational research aims to reveal where and how current forms of schooling are non-educational, and tend to create the changes necessary to perform schooling - to make it more

> educational. This is the core idea of improvement. (Kemmis & McTaggart, 1990, page 36)

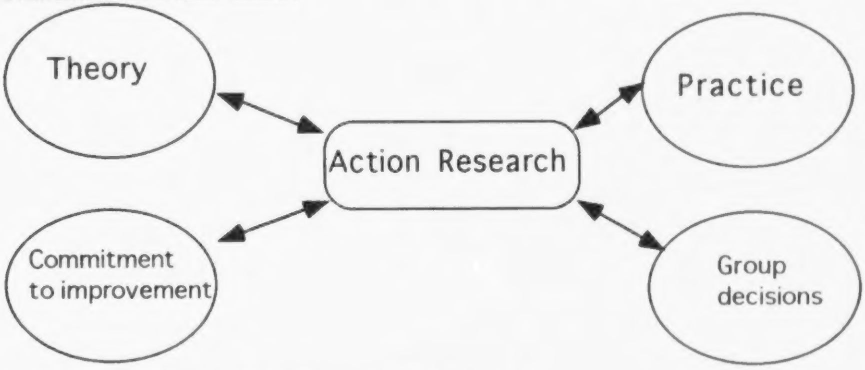
The selection of the action research orientation

Action research is a form of collective self-enquiry undertaken by participants to improve the rationality and justice of their own social or educational practice. John Elliot indicates:

> The fundamental aim of action research is to improve practice rather than to produce knowledge. The production and utilization of knowledge is subordinate to, and conditioned by, this fundamental aim. (Elliot, 1991, page 49)

To be action research, the approach must be collaborative, though it is important to realize that the action research of the group must be achieved critically examining the actions of individual group members. The critical feature highlights the terms *action* and *research*, which means: trying out the ideas in practice as a means of improving the curriculum. Action research links theory with practice.

Figure #2 Kurt Lewin's view of action research:



Taken from (Kemmis & McTaggart, 1990, page 11).

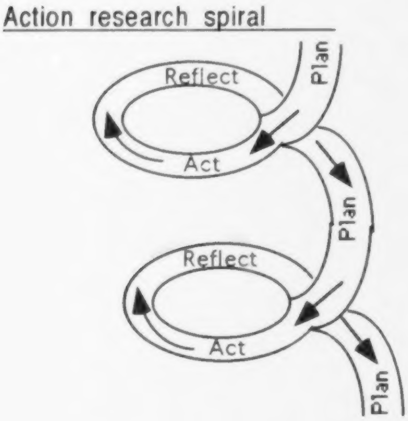
A distinctive feature of the research is that those affected by the planned changes have the primary responsibility for deciding on the course of critically informed action which seems likely to lead to improvement and for evaluating the results of the strategies attempted.

Describing the action research orientation

Action research can be described as proceeding in spiral steps. However, action research should not be consider a procedure. Elliot cautions against the danger of "interpreting methodology as a set of mechanical procedures and standardized techniques rather than a cluster of dynamic ideas and principles which structure, but do not determine, the search for understanding within a pedagogic process" (Elliot, 1991, p. 1). The four steps are listed below:

1. The group identifies a problem or thematic concern.
2. The group decides to work together on the thematic concern.
3. The idea prompts a reconnaissance of circumstances of the field, and fact finding about them.
4. The group decides on a field and makes a preliminary reconnaissance. This leads to a general plan for action.

Figure #3: Elliot's action research spiral


Action research spiral

The plan

The plan follows a spiral of action, observation, reflection, and revision. The plan is constructive and by definition must be prospective to action; it must be forward looking. The plan must recognize that all social action is unpredictable and therefore risky. The plan must be flexible enough to be responsive to modification.

The critical orientation, accommodated by the action research approach, is premised on the idea that life can be improved., in opposition is the correspondence theory, which suggests that the process of schooling is merely a reflection of socioeconomic class structure, and relatively impotent in terms of changing society. The critical orientation is founded upon contestation theories of social reproduction, which advocate that social change is possible through education. Contestation theory views humans as agents capable of effecting change in the process of constructing the social reality of education and its relationship to wider society (Kemmis & McTaggart, 1990).

The critical orientation is founded on the premise that life can be improved, whereas the situational-interpretative orientation is founded on the premise that life is a mystery, and that reality is inter-subjectively constituted. The empirical-analytical orientation is founded on the premise that life is predictable and can be explained with certainty. Reality, according to the empirical-analytical orientation, exists apart from the subject (Aoki, 1978). One of the benefits of action research is derived from a world-view that considers the interaction of self-and-world. Humans transform and reflect upon their world to create reality. Aoki (1978) describes reality as being created through praxis (thought and action). The researcher is not separated from action. Teachers do not step out of their pedagogic role to become researchers. Practitioner reflection and action are simply two aspects of a single process. John Elliot indicates that "all too often the idea that academic inquiry constitutes a form of teaching, and vice versa, gets lost." (Elliot, 1991, page, 14). Elliot emphasizes the unifying goal of action research:

> Action research improves practice by developing the practitioner's capacity for discrimination and judgment in particular complex, human situations. It unifies inquiry, the improvement of performance and the development of persons in their professional role. (Elliot, 1991, p. 52)

My interest in teachers' use of practical knowledge for decision-making is also supported by the action research approach. The improvement of performance and the development of persons in the professional role, as described in the quotation above, is supported by what Elliot calls practical wisdom. Practical wisdom is described as "the capacity to discern the right course of action when confronted with particular, complex and problematic states of affair" (Elliot, 1991, p. 54). Practical wisdom, like practical knowledge, informs judgment.

What action research is not:

- It is not a discussion about teaching problems; action research is more systematic and collaborative in collecting evidence.
- It is not simply problem solving. Action research is also problem posing.
- It is not research done on other people. Action research is a collaborative venture.
- It is not a scientific method applied to education. It adopts a social sciences methodology which concerns itself with the researcher as an agent of change (Kemmis & McTaggart, 1990).

Collaborative professional development through action research

Action research, by its very design, can have an impact on the learning in classrooms during its course rather than having to wait until the research results are analyzed to be translated into classroom practice (Casonova, 1989). Eisner (Eisner, 1993) points out that all too often schools imported concepts and theories from other fields and attempted to translate them into classroom practices. The assumption that a plan is transferable as a methodology because it has proven successful in one environment has a long history of failure in education (Herman, 1992b).

Because action research arises from a commonly perceived problem or reason for change, the expectation of a pre-prepared method for bringing about the change is not evident. Teachers are not viewed as the implementors of an external plan, derived for the "generic" classroom, but as innovators and initiators of change. Teacher and researcher work collaboratively in seeking a direction for change. Together they formulate a plan that is context specific. Teachers are not evaluated on the basis of how successfully they were able to carry out someone else's plan. They are co-contributors to a plan that both teacher and researcher anticipate will have to be modified. Ideally, students are included in formulating the plan. The goal should be of benefit to all who either directly or indirectly participate in the action research project. The plan does not only improve classroom practice but it also serves as a vehicle for informing teachers and researchers about classroom practice. The primary responsibility of the researcher is in providing a systematic documentation of the changes and a focus for inquiry on which reflective practice can be founded.

Allan and Miller (1990) outline four benefits for teachers if a collaborative environment is adopted:

1. Because teachers ask questions based on their own classroom needs, they become the originators of innovation. Ownership is assured because they have helped identify the problem and provide direction for change.
2. Teachers confirm their teaching ability and its effects upon students by careful collection, analysis, and interpretation of their data.
3. Within the collaborative setting teachers and researchers have the opportunity to gain support from other professionals and to share their classroom experiences with their peers.
4. The collaborative research group can support the teachers in a new professional role of presenters of their knowledge to unknown audiences.

Shumsky (1990) identifies the following benefits for collaborative action research:

1. Action research provides mutual support for teacher and educator. Teachers learn that many of their classroom frustrations are shared by colleagues. Researchers have the opportunity to view their interpretation through many windows. The context for change is made pivotal to the research rather than a variable that the researcher seeks to remove.
2. Cooperation in action research releases creativity and critical thinking. Discussion with colleagues opens alternative possibilities.
3. Cooperation in action research promotes change. The group, by providing support, helps promote risk-taking. A cornerstone to enacting a new plan is the criticism of existing structures. Support is required to overcome long held traditions.
4. Cooperation in action research promotes consensus. The researcher does not manipulate the group to a desired end but follows the spiral of plan formulation

and reconnaissance. Individuals within the group look for affirmation and support in new endeavors.

Finding an action research theme

The *Action Research Planner* (Kemmis, 1990) provides an activity that can assist in uncovering a theme for the action research project. The four basic categories Schwab (1969) regarded as the commonplaces of education: (teachers, students, subject-matter, and milieu) are used to construct a table of inventions.

- Teacher
- Students
- Subject matter
- Milieu

According to Schwab (1969), all educational situations involve the interactions of the four components. The table of inventions is derived from ancient Greece -- as a way of organizing thought for public speaking. The systematic structure provided a structure for reviewing and discussing a topic or identifying a theme. The method or way of working was known by the Greeks as the *"inventio"*. A table of inventions can be found in appendix 9.

Collaborative change

A plethora of curriculum initiatives over the past two decades have demanded that teachers give up much of their own orientations, styles, and beliefs (Perrone, 1978). The imposition of "teacher proof" curriculum, designed to alleviate difficulties experienced by the weakest members of the profession, has alienated creative teachers by forcing them to implement the "master plan". Effective change requires a collaborative approach in which the creativity of professional educators is maximized by a supportive environment that encourages risk-taking.

The work of Baird and Mitchell (1986) indicates that student attitudes and acceptance of new teaching strategies are greatly influenced by the attitudes of the teacher. This indicates the importance of shared ownership of teachers, administrators and school community. Nowhere does this principle become more important than working on a model to provide alternative strategies for student assessment. An Australian action research study (Tobin, et al., 1988) suggests that teachers will only adopt new strategies once they value the strategy and develop a high skill level in the strategy. This suggests that teachers need time to develop and modify these strategies. Encouragement is a most important feature. Teachers, administrators, and students need to know that some strategies will fail altogether or be reworked until they provide the desired results. Cooperative research orientations, in contrast, permit the introduction of an external plan.

Teachers can be asked to accept the external plan. Goodlad (1983) explains that for many years, teachers have relied almost exclusively on commercially prepared materials to provide instruction. Accepting external support resources, without attempting to adapt them to the learning context of the school, has created a depersonalized world of learning for both teacher and student. Effective change occurs when policy setting and planning occur within the group responsible for implementation (Allan, 1990).

Change from within the school

> Outcomes-based assessment and instruction will eventually require altering the basic premise that testing must come from outside the school and be used to check up on students and teachers (Pipho, 1989, p. 662)

A critique of the 'top down' approach to curriculum development and assessment is well documented (Aoki, 1978; Aoki, 1984; Baird, 1986; Corey, 1990; Darling-Hammond, 1991; Eisner, 1993). The purpose of the action research group was not to produce a standardized set of assessment templates, appropriate for all teachers in all situations, nor was it designed to produce a teacher-proof approach for the organization of portfolios. It merely served as a forum for sharing ideas, critiquing approaches, and interpreting success or failure. Teachers continuously personalized assessment templates, examined other indicators for performance-based-assessment, and returned to more traditional approaches for assessment when they felt uncomfortable. During an interview Marvin commented:

> I don't like being given a device for marking and being told to use it. When it doesn't work, the only one to blame is you. It must be the way you used it. If you develop the assessment tool, you know that you may have to modify it. Those developed outside of the school are sold as if they are pre-tested. The fact that everyone else has used it successfully indicates that something is wrong with you.

Traditional forms of assessment were not abandoned by teachers in the action research group. Formal exams, very often in partial multiple choice format continued to be the most dominant form of assessment. The change in some classrooms was pronounced, while in others it was much more subtle. Not everyone in the school needs to do the same thing. Thad commented:

> I don't like being forced into a mold, even by the majority. There should be some opportunities to draw on individual strengths. I guess that is also true about students. Maybe all of them don't have to be assessed on the same things.

Later in the same interview, Thad said:

> I do believe that it is important for teachers to be aware of what their colleagues are doing. Even if you're not going to do the same thing. Students always know what is going on in other classes. They may want to know why some things are different.

The school must afford time and recognition for those involved in the change. The experience at two of the research sites demonstrated the importance of support from school administration, colleagues, and parents. A representative from one of the action research teams provided a presentation to the school advisory council and received support for the new venture. A professional development day was also set aside and substitute teachers hired for a one day discussion. A teacher who asked not to be identified for this comment indicated that, at his school, the administration did not oppose their work on assessment but that he felt that he worked in isolation. He went on to remark that:

> I never felt any opposition. I believe that (the administrator) trusts our judgment, but he really only wanted to know if it would hurt our diploma marks.

Origins of the action research groups

I instructed two courses for the Department of Secondary Education, under the auspices of Special Sessions in spring 92 and summer 92, addressing the STS approach advocated by new science programs in Alberta. The focus of the course work for both courses was the science 10 program implemented in September 92. Teachers who enrolled in the courses expressed concerns about teaching unfamiliar content. Initially, the course was viewed by most teachers as an opportunity to regain long-lost information in biology, chemistry or physics. For many science teachers who had concentrated on the physical sciences, the course provided an opportunity to learn some biology. Similarly, hesitant biology teachers acknowledged the need to learn some physics.

The spring session course provided an informal beginning for action research. Thirty-five teachers, the entire science departments from the three different St. Albert high schools, attended the course on four weekends in March and April of 92. Friday night seminars were followed by a full day workshop session on Saturday. Attending class in their own setting was, I believe, a significant feature of the course. Activities were carried out in their own facility, and therefore, contextualized within their teaching environment. Restrictions placed on activities associated with a lack of equipment, limited classroom space, and time restrictions, due to time tabling, were addressed. Because the participants involved in the course continued to be active classroom teachers, their orientation toward curriculum change and reflection about topics presented in the course were inextricably linked with their own classroom experiences.

The special session course provided an excellent opportunity to view science teaching and learning differently. My role as facilitator was assured, in part, due to my lack of expertise in chemistry and physics. I was not seen as the content expert. A collaborative atmosphere was achieved by assigning expert groups to develop an activity from the science 10 program in a manner that reflected one of the three primary themes: nature of science, science and technology, and the societal aspects of science and technology. A course that began as a tutorial for assimilating new facts evolved into seminars and workshops that focused on pedagogy. The 35 hour course provided a significant start that enabled teachers to establish a common language for discussion and the comfort required to work with each other. The shifting focus for the course brought forth a second advantage of the classroom structure, one I would like to assign to careful planning but occurred by happenstance. Because the teachers enrolled in the course were actively involved in teaching, discussions leading to new teaching strategies were often introduced the following week. Many of the teachers in the course, interested in student's perceptions of what they saw as new teaching strategies, elicited reactions from students and colleagues. One teacher recorded in his journal;

> I began talking with my students about their experiences and this changed not only my teaching but their perceptions about my teaching. I am sure many of the students have never been asked if a teaching strategy worked. They were surprised that I didn't have all of the answers. I think the students no longer thought of teaching as something I imposed upon them.

Teacher and student perception about the teaching strategies could be discussed among colleagues. Discussions about similar experiences were shared and in many cases compared during informal discussions.

By far the most lively and perplexing discussions revolved around assessment strategies. Although many teachers willingly tried different ways of presenting laboratory activities, and embraced the idea of using case studies, role-playing scenarios, and interactive debates, they expressed concern about how these new strategies would be evaluated. What would happen when the cloak of scientific objectivity was dropped? Would their teaching and assessment be viewed as less objective? Would students raise concerns about fair treatment?

The course in St. Albert ended with a call for further work to develop assessment strategies. A series of follow-up workshops and informal meetings during the later months of 1992 and early 1993 formalized research questions. Subsequent work with 7 student teachers placed at the three high schools, as a university faculty consultant, during February, March, and April of 1993, allowed the action research project to continue.

Many of the student teachers had used portfolio assessment in their EDSEC 200 course and, therefore, were able to provide an added dimension to the research. Although novices to teaching, the student teachers were comparative experts in new assessment techniques. The arrival of student teachers marked the beginning of a site-based action research group. A single high school, identified as Northern High School, was identified for continuing research toward my doctoral dissertation.

During February of 1993, the project was expanded to two Edmonton Catholic High Schools, identified as Eastern and Southern High Schools. Teachers at both schools expressed an interest in exploring an alternative method for student assessment. A series of school-based meetings were set up to discuss assessment alternatives. Although some of the experiences of other teachers were shared, the group was encouraged to focus on things that took place at their school. The action research groups at the three high schools identified by the study worked independently.

The action research group at Northern High School was not used as a model for action research at the other two schools, although aspects of its success were discussed. The objective for working with teachers in two other high schools was not an attempt to replicate a methodology developed at Northern. The very idea of developing a method that can be validated by reproducible results stands in opposition to the philosophical orientation of the research. No claims of reproducibility will be made. At best the two additional sites will provide the reader with a wider panorama of divergent ecologies as action research groups attempt to link student's assessment with their modes of learning.

Summary

Action research is not a methodology to ensure change, but an orientation in which researcher and teacher can work collaboratively toward a common goal. For my study, the research approach did not rise from a pre-determined plan, but rather evolved from a series of events that channeled me toward collaborative work.

The advantages of action research can best be viewed by the development of processes used to bring about change. In this study, I shared many of the same problems experienced by colleagues. Because I was an active participant, rather than a detached observer or recorder of teacher practices, I developed greater empathy for teacher concerns and I was afforded greater candor from teachers during discussions.

The dual role of participant observer also created a variety of problems. In my journal entries I often queried whether my observations were distorted by my desire to create change. The commonality of failures and successes that I shared with teachers during the project also created an image of me having greater expertise than I actually had. Some teachers believed that

I truly had a definitive vision of the end of the journey. Something I still do not possess. I constantly battled the seduction of assuming the role of expert. Although I wanted the changes to be clearly those of my colleagues, I did intervene on a number of occasions to provide my personal suggestions. My role as active participant, capable of sharing ideas within the action research group, and the image of expert, sharing his knowledge and providing guidance, were often blurred.

Chapter 4: Evolution within the Action Research Groups

Overview

In this chapter I will provide the reader with a closer look at each of the action research groups and the evolution of the project at each site. Events unique to each site will be explored, to provide a context for understanding the divergent concerns and approaches to assessment. The description of the action research project is organized into the three phases as identified by Lewin (Elliot, 1990); the initiation phase, the development phase, and the reconnaissance phase. According to Lewin's model, the phases progress as a spiral of planning, action, and reflection. According to his model, the resolution of one problem is followed by the creation of other problems. Such a model suggests a chronology of linked events; however, that is not what I found in my project. In most cases, one concern did not progress to a resolution before another concern surfaced. Most often, different problems emerged simultaneously. A plan for one problem did little or often nothing to solve the other problem. A chronology of events was not possible because a single spiral was never found. Multiple spirals could be identified. Some spirals provided a fast-acting plan, action, and reflection, while others took much longer. Many times it was difficult to recognize when the problem was identified or the plan instituted.

The grouping of events into the phases (initiation, development, and reconnaissance) is intended to present my findings in an organized fashion. Because each site is distinctive, I have purposely provided three different summaries for each of the research groups. Keeping the summaries separate has also helped me guard against the temptation to compare each of the sites. The seductiveness of developing cause-and-effect relationships by comparing different sites was never my intention. I will avoid the trap of rationalism by providing and analyzing why aspects of authentic assessment worked better in one location than another. I leave comparisons to the reader.

Reflection upon the action research conducted at Northern school

As indicated earlier, Northern high school was one of the schools involved in the pilot project. Following the pilot study, the action research group at Northern high school continued and became part of my major study. A number of unique characteristics can be identified at Northern high school. The influences are grouped under three main headings: central office support, university course structure, and contributions by student teachers.

Central office and administrative support

- The research project began as a cooperative venture between the University of Alberta and St. Albert Protestant School Board. The need for teacher inservice had been initiated by Robert Hogg, then Director of Curriculum Development, in consultation with science department head at both high schools. My position as course instructor affected the relationship between researcher and teachers.

- Seminar money was allocated for follow-up workshops and discussions. Substitute money was allocated to free teachers from the classroom to work on assessment templates.
- Selected articles were photocopied and a bibliography was prepared on alternative assessment and given to teachers to encourage further reading.
- Assessment questionnaires were prepared, collected and analyzed following the university course, and workshops were conducted to determine whether teachers' need had been addressed. In addition, follow-up meetings were set with department heads and me to ensure that the action research project provided something of value to the teachers involved.

University course structure

- Because all members of the science department had registered in a University course, a shared vocabulary developed. The research question, although identified through discussion by the university class, was formalized by the researcher.
- The university course also provided a forum for sharing ideas, proposing solutions, and reporting on initial findings.
- The initial class also provided a significant amount of meeting time in a concentrated time frame.
- The focus for changes in assessment was directed toward science 10, a course that the majority of teachers at Northern High School believed that they would eventually teach. Physics, biology, and chemistry teachers were able to direct their discussions at common activities. At the two other schools, content differences affected the level of assessment discussion.

Contribution by student teachers

- The infusion of three student teachers, who concentrated primarily on science 10, facilitated the action research. Each of the student teachers worked with two or more cooperating teachers within the department. This arrangement permitted the cooperating teachers an opportunity to speak with colleagues about the direction of the proposed changes and view many of the changes as implemented by the student teacher.
- The student teachers were also perceived to have some expertise in using alternative assessment. Their ideas were often sought-out and most often valued.

Organization of the action research group

- All teachers within the science department were invited to the meetings. On most occasions most teachers were able to attend the after school meetings.
- No chairperson was ever assigned for the meetings, although teachers within Northern High School looked to Henry, the department head, as both their formal and informal leader.
- Initially, student teachers brought items of concern to the meetings. However, as the meetings progressed, teachers interjected scenarios, classroom success stories, and questions that troubled them about using alternative assessment. After the first meeting Henry and I shared an activity and assessment strategy that we had tried with our classes (see appendix 9F). Some sharing continued following that meeting; however, a number of teachers in the research group felt more comfortable using the assessment ideas of others and modifying them for their classes, rather than developing their own.
- The meetings were conducted in an informal setting and donuts and coffee were supplied by the science department.

The following observations have been organized into three sections, the initiation phase, the development phase, and reconnaissance phase. During the initiation phase, the problem was identified and concerns were identified. During the development phase, the plan was structured to meet the concerns and implemented. During the reconnaissance phase, the plan was formally evaluated. Although three separate phases have been identified for the purposes of presentation of findings, it is important to note that the phases did not occur in linear sequence. In most situations a single concern leads to a plan and then action. Most often, the action raised other concerns, which, in turn, led to a restructuring of the plan.

A. Initiation phases

1. Concern: increased marking

Teachers identified that the new science 10 course, because of an increased emphasis on laboratory activities, would require a restructuring of current assessment practices. Because each component of the laboratory activity was marked the day after students completed the activity, teachers were concerned that moving toward an activity-based course would greatly affect their marking load. Approximately 10 minutes were required to mark each laboratory activity. A concern for the added time required to prepare materials for the new course in addition to increased marking time was also expressed. Teachers also indicated a general degree of dissatisfaction with the text that had been adopted for the science 10 curriculum and indicated that increased planning time, to structure a worthwhile course for science 10, meant that marking time must be decreased.

Response : a plan was struck to begin trying portfolio assessment in science 10 classes. Implementation was considered optional for semester one but the entire department would begin portfolio assessment for 1993.

2. Concern: assessment strategies for STS

A number of the teachers also expressed concerns about the introduction of social issues into the curriculum, under the direction of STS (science-technology-society). Although only a few teachers expressed concerns about the infusion of social issues and a technological way of thinking with commonly held content, some concerns were expressed for how these components would be assessed. Concerns were expressed about having students complete activities that would not be reflected in a report card mark. Yet many of these activities were difficult to transfer into a number for report card purposes.

Response: a social issues template was developed, used, modified, and personalized by a number of teachers. See appendix 6C. Not all teachers use the assessment template.

3. Concern: Acceptance by parents and students

Initial concerns about accountability were raised by teachers. Would parents and students accept different ways of marking that moved from a highly directive marking guide to one that was more holistic. A number of concerns about the reliability of holistic marking were raised. Teachers had been modeling marking guides, from those used on diploma exams, that identified the value of each question within a laboratory activity. Because each question had a specific answer, the totals could be added and an aggregate reported for the laboratory. Teachers did not believe that standardized assessment practices by the Student Evaluation needed to be changed, but that their new practices would be different than those used in standardized assessment.

Response: Teachers acknowledged that they must live with the doubt.

4. Concern: Preparation for diploma exams as a primary responsibility.

A number of teachers expressed the concern that what we do in classrooms at the grade 10 and grade 11 levels should mirror what happens at grade 12. These teachers did not argue for an acceptance of this analytical marking scheme because it was superior or even more appropriate for their students. They believed that any familiarity with the marking scheme would provide students with an advantage during their diploma exams. During one of the seminars a teacher with twenty-seven years of experience expressed that he felt his primary responsibility as an educator was to ensure that students gained every advantage possible when they wrote their diploma exams.

Response: For most teachers the concern has not been resolved.

5. Concern: New curriculum orientation requires different assessment strategies.

The action research group believed that current assessment practices, because they were directed at finding the "correct" answer, failed to provide any way of assessing other components identified by the new science curriculum. The ability to use technological thinking to alter laboratory designs and rework problems, the ability to work cooperatively in groups, and the ability to recognize social issues and alternate points of view from emerging science and technology were identified as important.

Response: The desire to have activities move away from a teacher-directed orientation, where students would follow a set procedure, to less structured experiments, resulted in teachers re-working of many of the activities found in different science courses. In newly developed laboratory activities, students were expected to modify some of the experimental procedures, create their own procedures, identify the problems to be solved, or assess procedures developed by other groups.

B. Development phases

1. Plan: Portfolios for science 10 (September 1992)

The idea of using portfolios, as representations of student's work, had been introduced during the university course. Advantages and disadvantages had been discussed within the course. Small group discussions and the reporting of findings to the large group was employed as a technique for inquiry. Following the course, a meeting was held in June of 1992 to identify components for inclusion within the portfolio for science 10. It was decided that three portfolios would be assigned each semester to correspond with reporting dates. Laboratory activity, assigned questions, and journal entries were selected as works that should be included as representative for student work. A combination of teacher and student choices for laboratory assignments was selected for the first portfolio. Teachers of science 10 did not have to use portfolio assessment, but they were strongly encouraged by Henry. Those who were most reluctant were somewhat resistant to teaching science 10.

2. Action: Initially, a common format was used for portfolios. Not every science 10 class used portfolios. Two strong leaders emerged in Northern School, the teachers identified as Henry and Albert. Both encouraged colleagues to try portfolios.

3. Action: Student teachers begin using portfolios in science 10 (February 1993). Cooperating teachers had the opportunity to view and monitor student reactions to portfolios. All science 10 classes began using portfolios at Northern High School. It is important to note that some of the cooperating teachers who did not teach science 10 did not begin using portfolios. This was especially true of part-time teachers within the department. Science 10 teachers who had used portfolio assessment during the first semester acted as informal guides and provided moral support for both the student teachers and the cooperating teachers who had just begun using portfolios. Student teachers introduced projects as a part of the portfolio in all but one of the classes.

4. Action: Portfolios extended (February 1993) for other subject areas portfolio use was extended for chemistry 20, physics 20 and biology 20 at Northern High School, although not every teacher decided to use the portfolio system. Three teachers began using portfolio assessment with their grade 12 classes. No attempt was made to pressure the few reluctant teachers who did not try portfolio assessment to begin doing so. Henry, the department head, expressed frustration when speaking about teachers who were slow to change. He also expressed a paradox in that he also wanted staff members to maintain their professional autonomy in setting assessment strategies. Concerns were raised at scheduled meetings about the science department using different assessment strategies. At this point, the level of interest and commitment for changing assessment strategies appeared to wane for those who had not made changes.

5. Plan: Performance-based-assessment (September 1993) Under the guidance of Henry, the action research group at Northern High school began exploring the use of performance-based-assessment. Specific laboratory activities were identified by the group and templates were assigned to interested teachers.

6. Plan: Social issues' assessment identified (January 1994) The action research group began looking at different ways in which social issues could be assessed.

C. Reconnaissance phases

1. Evaluation: Teacher perceptions collected
Teacher interviews were conducted to determine the ongoing success of the plan. Bimonthly planning sessions were organized for 4:00 PM , the second and fourth Tuesdays of the month at Northern school. Most often eight teachers from Northern High School attended. Student acceptance, administration support, and grading concerns were discussed. Weekly meetings were conducted with student teachers, with an invitation to cooperating teachers from February 1993 to April 1993. All meeting were optional, but well attended.

2. Restructuring: Changing meeting time.
Action research meetings were reduced from 60 minutes to 40 minutes after the first meeting. No meeting were conducted for May, June, or December. Once my involvement with the action research team ceased in May 1993, meeting times were rescheduled to lunch time or department meetings.
3. Evaluation: Sharing results with other schools and colleagues.
A professional development day was organized in which teachers who had been involved in alternative assessment shared their experiences with teachers from two different high schools and teachers in other departments at Northern High school. Portfolio assessment strategies were scrutinized and experiences were clarified.
4. Restructuring: Many different assessment templates were developed for portfolios. The changes were most often a reflection of different course emphasis, activities completed, or reflected different projects conducted in science classes. For example, different laboratories were identified as compulsory by different teachers.
5. Restructuring: The description of marking schemes was changed to make it easier to read. Teachers found that they had to interpret many of the indicators used to describe scientific or communication skills for students.

Summary observations and interpretations

- During the university course, the hosting high school would provide a Saturday lunch for all thirty-five teachers. The lunches, which progressively became more elaborate, served as a good-natured challenge between schools. The tradition continued for the 40 minute meetings of the action research group. The social atmosphere and relaxed format for sharing were an important part of the continued success of the action research group. The importance of developing social aspects for the action research group is supported by Sanger (1990).

- Allan and Henry emerged as both formal and informal group leaders, likely because they were most willing to take chances. My role within the group, despite its beginning as course instructor, evolved to that of an observer, and one who shared personal experiences from the other two action research sites.
- Despite a common starting point, tremendous variations were noted. Some teachers within the group embraced change and were willing to take a number of chances. They expressed few concerns about parent, student, or administration support. They did not see it as a problem because they perceived that the support had always been there. Teachers more reluctant to change expressed either personal stories of nonsupport or related stories that they had heard of, or people who they knew who had been defeated attempting to bring about changes.
- Although community acceptance was initially discussed as a potential problem, no parent, administrator, or central office personnel concerns had been voiced. I believe the communication established by Henry, the department head, was instrumental in avoiding problems.
- The re-thinking of assessment strategies caused a re-working of laboratory activities to create greater student independence. A more open-ended laboratory approach was adopted.

Reflection upon the action research conducted at Eastern School

I acted as both teacher and researcher at Eastern high school. I was recruited by the principal following the completion of my graduate course work in 1992. Although I had met many of the teachers at Eastern high school, I had never taught at the school. My assignment came at a somewhat sensitive time of restructuring and the fact that I was a personal friend of the principal meant that any action research project at this school would require some sensitivity. I spoke with three of the four teachers in the science department in October of 1992 and described some of the work that had been started at Northern high school eight months prior. I began teaching at Eastern high school in February 1993 and began using portfolio assessment in chemistry 20, biology 20 and biology 30 classes. My previous experience with portfolio assessment, as a teacher, had been at the university level. No attempt was made to ask any of the teachers to begin using portfolios but assessment strategies were shared and samples of student's work were reviewed.

In September 1993, three new science teachers joined the staff at Eastern high school to meet increased enrollment. For two of the three teachers, Jane and Larry, this was their first teaching position. Gordon had taught for another school board but was new to the system. I was asked by the two new teachers to the school to act as a mentor, for a program instituted by

Central Administration of the Edmonton Catholic School Board. Gordon, because he had previous experience as a teacher, was not required to take part in the program. The list below identifies contextual elements of school life that had an impact on the action research group.

Changes in school structure:

- The math-science department was combined with career and technology in June of 1993. Thad, the long time department head, resigned in June 1993. His resignation was not in protest to the changes. He indicated that he resigned to spend more time with his growing family. I was asked to take his place.
- At the end of September 1993, I was appointed acting assistant principal of Eastern high school for the remainder of the 1993-94 school year. Marvin was appointed department head.
- June 1994, the department was again restructured and mathematics, science, social studies, and English were combined to form a large department of academic studies. I assumed the role of department head of academic studies for the 1994-95 school year.
- Due to budget cuts, Gordon and Larry were re-assigned to other schools at the end of the 1994 school year. They continued to use assessment strategies employed at Eastern high school at their new schools.

Central office and administrative support

- The research project received permission and encouragement from Muriel Dunnigan, then Director of Curriculum, for Edmonton Catholic Schools. On three separate occasions, between September 1993 and 1995, the area Central Administration found money for science equipment that would allow Eastern high school to implement changes in both teaching and assessment. The total sum of money was approximately $33 000.00.
- The school administration at Eastern high school, despite changes in personnel, was also highly supportive. Money and support were received to help integrate science with a new career and technology studies laboratory being constructed within the school September 1993. Additional science equipment, to a sum of approximately $25 000.00 was secured from a grant. The science budget also received a substantial increase in funding from the school.
- Substitute money was allocated to allow teachers to initiate discussions about alternative assessment twice throughout the research project. Two full-day sessions took place September 1993 and October 1994. In addition professional development money was secured so that all members of the science department could attend the ATA Science Council Conferences in October 1993 and 1994. Professional development money was

also allocated for regional chemistry, biology and physics council workshops in May 1994 and 1995. Only one of the teachers in the school had ever attended an ATA Science Conference prior to 1993.

- In September 1993 a full-time laboratory technician had been secured to improve the delivery of science programs. Prior to that time, only a 0.4 time had been allocated. The experienced laboratory technician was pivotal in encouraging teachers to begin doing more laboratory work. Prior to her arrival very few laboratory experiments were conducted at the school. An average of 93 labs were prepared each month for the 1994. 95 school year.

Contributions by students.

- Because some students became involved in the discussions about alternative assessment, they also served to convey their impressions about alternative forms of assessment to other teachers. Students became agents of change.

Organization of the action research group

- All teachers within the science department were invited to the meetings. On most occasions most teachers were able to attend the after school meetings. One teacher never attended a single meeting. He was not included in the action research group.
- No chairperson was ever assigned for the meetings, although teachers within Eastern High School looked to me to provide leadership and often expertise. I consciously attempted to dispel the notion of expert.
- Common planning for science 10 became an initial focus for developing alternative assessment strategies in a collaborative setting.
- The meetings were conducted in an informal setting and refreshments were supplied by the science department.

A. Initiation phases

1. Concern: increased time needed for preparation and marking.

The introduction of science 10 , in the fall of 1992, was not met with success at Eastern high school. Difficulties were traced to three main factors. First, the activity-based course required a different teaching style than what had been used in the past. A lecture, teacher-directed teaching style had been used by most teachers. In large part, the lecture-based approach was necessitated because of a dearth of science equipment. Second, the support for preparing solutions, organizing laboratory stations, and assisting students did not exist. Although a part-time laboratory technician had been employed, she lacked science background. Most of her tasks involved clerical and typing duties. Third, teachers had taught only one discipline for more than the past ten years. Of the four full-time and one part time sciences teachers in the department, only Marvin taught two subjects (chemistry and biology). A strong dissatisfaction with the textbook accompanied with the increased time required to prepare materials for the new course and increased marking time because of the movement toward more laboratory work made the staff receptive to exploring new avenues for assessment.

Response : Plans to begin portfolio assessment for September 1993 for any science classes where teachers were interested. All science 10 classes adopted portfolio assessment.

2. Concern: rapid changes within the science department.

The introduction of four new teachers to the department changed interactions within the department very quickly. Laboratory rooms that sat free were most often booked by 1993. The introduction of new teachers also helped encourage change of assessment. The new teachers were eager to try alternative assessment strategies. Three changes of department heads between June 1993 and September 1994 helped erase some long held traditions, but also created some disequilibrium. Prior to my arrival in February 1993, the last new teacher that had been added to the science department had occurred 12 years prior. Two of the four members of the science department had spent their entire teaching career in the same school (both teachers have more than 25 year experience).

Response: No changes were to be mandated. Teachers who had opted to try some alternative teaching and assessment strategies were given support, but the support was also provided for those who wanted to view what was happening at arms length. New teachers to the department were sensitive to the traditions and approaches of those long-serving members.

3. Concern: Manipulating change

In my journals, I expressed the concern about being in a position to mandate change. Although I felt that I could serve as an agent and encourager of change, I didn't want to orchestrate change for the sake of my research.

Response: I took a low key approach to working with the long serving members of the department. Because new teachers to the department wanted to work together, they provided an opportunity to begin doing action research. Initially, the action research group arose from common planning for science 10. The planning sessions were informal and often very social. Eventually, planning sessions were attended by other members of the science department.

4. Concern: Acceptance by parents and students
Like Northern high school, initial concerns about accountability were raised by teachers. Would parents and students accept different ways of marking that moved from a highly directive marking guide to one that was more holistic? A number of concerns about the reliability of holistic marking were raised. Teachers had been modeling marking guides, from those used on diploma exams, that identified the value of each question within a laboratory activity.
Response: No concerns from parents were forthcoming. Teachers decided to continue working with some degree of uncertainty.

5. Concern: Diploma exams dictate accountability
The concern of using a grading system and approach to assessment that was not reflected on the diploma exams was raised. Because diploma exam results had been slightly below the provincial average in 1992, teachers had received a letter from Jim Dinning, then Minister of education, pointing out that their marks were below provincial average. The letter was referred to often at meetings. Especially for teachers who invested a large part of their career at the school, the letter served as a reminder of what really counts. The value of increasing the manner in which we value laboratory experiences, the interaction of science and society, or technological thinking became blurred when the most important part of assessment focused on pencil and paper exams. The idea of moving away from worksheets, the long-held approach in favor of debates, projects, and laboratory activities provided, the direction advocated by the new curriculum reform, created a paradox for teachers. How could they satisfy both masters, when, they believed, curriculum and evaluation directions are incongruent?
Response: We have identified the dichotomy of contrasting expectations and recognize them as part of the turmoil created by mixing educational practices with a political agenda. They continue to be a source of frustration.

B. Development phases

1. Plan: I began portfolio assessment in February 1993, and shared examples of student work and my marking guide throughout the semester. The plan was shared with students in my classes and an initial class was used to explain the purpose of portfolios.

Action: Student reactions and impressions were sought as portfolio assessment began. Students were asked to make weekly journal entries, briefly writing about their impressions of selecting work for their portfolios. Six grade 11 students were identified for in-depth interviews.

2. Plan: Increased teacher support to allow for changes

Action: Additional equipment and an experienced full time laboratory aide allowed teachers to diversify the type of assignment that students were given. This permitted the greater infusion of laboratory activities into their lecture-based approach.

3. Plan: Optional use of portfolios for science classes (September, 1994)

Although implementation was considered optional, all science 10 teachers in the department agreed to begin portfolio assessment for September 1994.

Action: I worked with three new teachers new to the department and Marvin to identify compulsory and optional activities for the portfolio. One major project was included for each of the three portfolios: a research assignment and debate on clear-cutting temperate rain forests, writing a murder mystery that could be solved with science clues (microscope unit), and using a computer simulation (*More of the Incredible Machine*™) as a Science Olympics. Those who were most reluctant were somewhat resistant to teaching science 10. After the first report card, all science 10 teachers began using portfolio assessment in all of their science courses.

Action: Student teachers use of portfolios in science 10 (February 1994)

Teachers had the opportunity to view and monitor student reactions to portfolios. Teachers who had used portfolio assessment during the first semester acted as informal guides and provided moral support for both the student teachers and the cooperating teachers who had just begun using portfolios. Portfolio assessment was used by all but two members of the science department.

3. Plan: Social issues assessment identified (September 1994)

The action research group began looking at different ways in which social issues could be assessed.

Action: An assessment template was developed. See appendix 6C.

4. Plan: A peer assessment template was required for the project developed for report card #2 in science 10. Students were to construct a murder mystery and give it to another group to solve. Each member of the group that constructed the murder mystery was required to assess group cooperation and then the murder mystery that they attempted to solve.

Action: See Appendix 7A and 7B for the outline to the murder mystery and appendix 6A and 6B for the assessment templates.

5. Plan: Performance-based-assessment (December 1994)
Specific laboratory activities were identified by the group in science 10 and biology 20. Templates were assigned to interested teachers and developed.
Action: An ongoing development of performance-based-assessment strategies built for specific labs. We also field tested five different laboratory activities for Alberta Education that provided performance based assessment in the areas of physics 20, biology 20, chemistry 20, and science 10 (June 1995).

C. Reconnaissance phases

1. Evaluation: Student perceptions sought (February 1993)
Student interviews were conducted to determine the ongoing success of the plan. Open-ended journal entries provided direction for formulating questions for the interviews.
2. Restructuring: Changing structure of portfolio grading sheet (February 1994)
Gordon reworked the grading sheet to change the columns in which marks were collected. The new grading sheet was accepted and used by the teachers involved.
4. Evaluation: Sharing results experiences (September 1993 to June 1994)
A professional development day was organized in which teachers who had been involved in alternative assessment shared their experiences with other teachers in the school.
5. Evaluation: Concerns about the reliability of portfolio marking were addressed by Gordon, Jane and myself. Each teacher identified five portfolios from their science 10 class. Because all three teachers did group planning all activities identified for the portfolio were identical. Each portfolio was graded on a scale from 1 to 20, on a separate page, by each of the teachers. Teachers were not permitted to see each other's grading. After all 15 of the portfolios had been graded by each of the teachers the marks were compared. Of the 15 portfolios marked by all three teachers (n= 45), only 18 had different scores. But most importantly, only four of the portfolios had a grade differ by more than one mark. In every case, the teacher who taught that student gave the higher grade.
6. Restructuring: Performance-based-assessment template used (September 1993 to June 1995).
Many different assessment templates were developed for performance-based assessment. We found that generic templates did not work well in this area.

Summary observations and interpretations

- Administrative support was essential in permitting the action research group to succeed. The influx of money into the department and the securing of a very competent, full-time laboratory technician helped teachers view a vast array of rapid changes as generally beneficial. I believe any change without some support, to help the teachers with their daily tasks, would have been ill-received.

- Many informal gatherings, such as a champagne toast when Larry, Gordon, and Jane received their first pay cheque, helped form a sense of community. The sense of community was especially important because all three aforementioned teachers received a notice from the school board in late March of 1994, indicating that their contracts would be terminated should the tenured teachers employed by the school board fail to settle their negotiations. Ironically, a device that many believed would divide non-tenured and tenured teachers helped serve to create stronger bonds. (The contract was settled and all teachers continued until the end of June).
- In part, the success of beginning change with less-experienced teachers could be traced to the high degree of competency evident by the new additions to the science staff. Not only did the new teachers display excellent teaching skills, but strong interpersonal skills in dealing with colleagues. They were sensitive about not posing a threat. No inter-staff struggles were observed in the science department.
- After the 1993-94 school year, set meeting times seemed more difficult to maintain; however, the project continued to grow during the 1994-95 school year. All but one of the experienced teachers, who had waited more than a year and a half, now began using portfolio assessment in all but diploma courses. A timetable for change can not be imposed.
- Because so many factors were at play during the school year, simple cause and effect relationships that identify reasons, for what the action research group believes is success, are not evident.

Reflection upon the action research conducted at Southern School

I had been a staff member and science department head at Southern High School from 1984 to 1990. Teachers at the school had been involved in various professional development activities including: exam development committees, curriculum committees, diploma marking, science conferences, student teaching programs, and inservice programs. In addition, I considered the teaching staff to contain many strong and committed teachers.

In February of 1992, I met with two teachers, Mary and Gary, who I believed had the enthusiasm and imagination necessary to derive new approaches for assessment. Because these two teachers were well respected by their colleagues, I believed that they would likely evolve as leaders for the school-based action research group. By design, the two teachers would try various strategies for alternative assessment and act as school liaisons by trying to encourage others to become involved. I explained some of the pilot work that had been done and Northern High School and some of the plans slated for Eastern High School. The philosophy for using portfolios and journals was discussed in two successive meetings in October and November 1992. Sample scoring rubrics developed at Northern High School for portfolios and sample student journal entries were provided.

Mary, the department head, and Gary, a veteran teacher, expressed interest in working to develop alternative assessment strategies that better reflected students learning and agreed to begin an action research group at the school. Although examples of authentic assessment could be shared among the three separate research groups, no attempt was ever made to work to common scoring rubric or produce a single methodology for initiating change. To ensure that change was initiated from within the school, each action research group remained autonomous.

Prior to embarking on using portfolios, Southern High School had been involved in an Alberta Education pilot to exam authentic laboratory assessment for skill development. Teachers had been given a pre-designed program and asked to field test various components for curriculum fit and classroom appropriateness. Two of the teachers expressed extreme frustration when talking about the pilot. Although they believed that many of the ideas were excellent, they indicated that the project proved to be extremely time consuming. Gary indicated an increase in marking time and clerical tasks to the detriment of planning and student conferencing.

Central office and administrative support

- The research project received permission and encouragement from Muriel Dunnigan, then Director of Curriculum, for Edmonton Catholic Schools. Support was received. Additional science equipment was secured from a grant but not linked to the action research project. A large portion of the money was identified for Laser Disks to help support the presentation of science information.
- A letter was written to the principal of Southern High School explaining the action research project. Once again support was received; however, unlike the other two high schools no additional money was requested or allocated for either teacher inservicing or professional development on assessment.
- In September 1994 the time allocated for a laboratory technician had been reduced from 1.0 to 0.6. Teachers indicated that the experienced laboratory technician was pivotal in allowing them to do laboratory work and supporting performance-based-assessment.
- Because science 10 was a full year course, teachers chose chemistry 20 as a single course to begin implementing alternative assessment strategies. Because Gary and Mary were the only chemistry 20 teachers, it ensured commonality within the program but, unfortunately, limited discussions with other members in the science department.

Organization of the action research group

- All teachers within the science department were invited to the meetings; however, only two teachers were to try alternative assessment. According to the plan established by the department head, in consultation with the principal, only selected teachers were encouraged to try new assessment strategies. Mary, the department head and Gary, indicated interest. Two other teachers provided with the option, decided not to become involved. Other teachers were to be encouraged to join the action research group for the second semester. Only Gary and Mary attended regular meetings. Both teachers were extremely enthusiastic and supportive.
- Mary acted as chairperson for the meetings. Her many and varied commitments throughout the school required numerous rescheduling of meetings.

Observations have been organized into two sections, the initiation phase, and the development phase. Because the project ended during the development phase, no reconnaissance phase has been presented.

A. Initiation phases

1. Concern: Increased time needed for preparation and marking. The Alberta Education pilot, on performance-based assessment, and the introduction of the new science 10 program underscored the increased time that was needed for marking. Although teachers in the school had for some time used activity-based programs, the increased emphasis on laboratories in the new science 10 further increased laboratory assessment and brought traditional marking time to a crisis. Teachers were receptive at examining alternative ways to better reflect student learning and manage the amount of time spent marking.
   Response : Plans to begin portfolio assessment for September 1993 for Chemistry 20 classes. Other teachers would be informed of the pilot and provided with materials, should they want to attempt portfolio assessment.
2. Concern: Implementing an external plan.
   Although common problems were discussed and the potential of the portfolios were presented, the two teachers at Southern High School adopted the outline for portfolio assessment being used at Eastern High School. The previous experience with the Alberta Education pilot, in which teachers implemented and externally derived plan, had a carry over in the development of portfolio assessment.
   Action: The pilot teachers field tested an externally derived plan. The project did not become a true action research project.

3. Concerns: The acceptance of change

Pilot teachers expressed great concerns about how the administration and community would accept change. Concerns about mark inflation and the disruption of a marking system that would no longer be reflected by diploma marks were expressed. Both pilot teachers expressed concerns about how the use of portfolios in one class might inflate marks thereby increasing class averages in certain subjects. What would happen if one chemistry class had higher marks than another because that teacher used portfolios? In no school was the importance of diploma marks stressed more. Both teachers expressed what they believed was the administrations expectation that diploma marks verify teacher marks. The practice of highlighting teacher assigned marks that varied more than 5% from the diploma marks was cited on four separate occasions.

Response : Portfolio assessment would not be used for diploma subjects and restricted to all chemistry 20 classes for February of 1993. Because Mary and Gary taught all chemistry 20 classes at the school, the potential for conflict could be reduced. No parent or administration concerns were encountered, although the teachers reported a number of student concerns with using portfolios.

B. Developmental phases

1. Plan : Setting of meeting times

Initial meetings were set for October and November of 1992. Subsequent meetings were established for February 1993, the established implementation date for beginning portfolio assessment. Because both Mary and Gary were involved in coaching, meeting times were difficult to set. A number of other science teachers in the school also had extensive extracurricular activities. In addition, Gary believed that the number of meetings at the school increased dramatically due to the school's involvement in aschool-based-budgeting pilot. Early weekly meetings were interspersed with a variety of sub-committee meetings.

Action: Meetings were difficult to establish. Rescheduling of meetings and the ideas of yet another meeting reduced communication. For the most part, Mary and Gary worked in isolation with only occasional opportunities for informal discussion. The plan was implemented without a true action research component.

2. Concern: Reduced teacher support

In 1993/94 the laboratory technician's time was reduced from 1.0 to 0.6. This made laboratory work more difficult to organize.

Response: Both pilot teachers speculated on the reduced emphasis on activities. This meant that students had fewer opportunities to chose activities for the portfolio.

3. Concern: Delaying student assessment with portfolios

Both Gary and Mary were concerned that students were not provided with immediate feedback. Student objections about delaying the assessment of assignments to three times throughout the semester. The teachers indicated that students felt uncomfortable not having every laboratory assignment evaluated immediately after it was completed.

Action: Portfolio assessment was abandoned after one semester.

4. Concern: The use of holistic scoring

Teachers expressed frustration at not being able to mark each component of the portfolio separately. The idea of incorporating journal entries into the portfolio was cited as one example. Mary raised the concern that students might not know if the journal entries are adequate, because the mark is subsumed into a mark for the entire portfolio. Excellent laboratory reports could be combined with poor quality journal entries and poorer quality chapter problems to give a distorted view of their progress.

Action: A more traditional scoring guide was used where individual questions were assigned a specific number of marks. Each assignment was marked in isolation after the first portfolio assignment was returned. Gary indicated that marking individual assignments increased marking time and that journals were very difficult to assess. After the first portfolio was returned, Mary decided to drop journal entries from the portfolios, while Gary continued to ask for them but decided not to assess them. Gary began using journals as open-ended letters to the teacher. He indicated that many students used the opportunity to use the journals as a means of informing the teacher about the quality of his instruction. Gary indicated that he appreciated knowing what things he did well and what things students found difficult to understand.

Summary of observations and interpretations

- Unlike Northern and Eastern High schools, limited administrative support existed at Southern High school. In large part this occurred, not because of any opposition to the project, but because the school's focus had been placed on the development of site-based-management. School budgeting and the development of a democratic plan for school decision-making consumed most of the attention. A great deal of in-school committee work had been directed toward these goals. The change in school management had a particular effect on Mary, the science department head, who became involved in a number of committees. No special days were set aside to discuss assessment practices or strategies for change. Teachers did not discuss the research project on assessment with administrators after a first initial meeting.

- The fact that only two of the members of the science department became involved in the study limited opportunities for discussion. Unlike the other two sites, discourse about change and modifications to the existing plan were very limited.
- A climate for action research was not established. Common problems were identified but an internal plan was not implemented. Examples from other schools served as a blueprint for changing assessment strategies. In many respects, the alternative assessment goals, although accepted, became yet another task that must be accomplished. In no other school was teacher focus so diverted and time so limited.

Chapter 5: Documenting changes in assessment strategies

Overview

In this chapter I will recount some of the problems teachers identified with current assessment practices. A common complaint was that traditional laboratory reports represented the only alternative to exams; however, some of the teachers felt that the reports assessed many of the same skills found on exams. Laboratory reports were seen as communicating the end-product of a students work without providing an indicators of what happened along the way. In the chapter I describe how laboratory assessment acted as a lightening rod for teachers' concerns. Increased laboratory emphasis, brought about by curriculum changes, elevated the level of solicitude.

A related theme within the chapter explores why teachers began using portfolio assessment. Although portfolio assessment resolved a number of the concerns that teachers had identified, portfolios created a whole new set of problems. Within the chapter, I describe the concerns teachers identified with portfolio assessment along with the problems it solved. Student perceptions are provided as a backdrop to understanding teacher concerns with portfolios.

Identification of problems with traditional laboratory reports.

Prior to using portfolios, teachers in the action research group expressed a common frustration. They believed a similarity of answers to laboratory reports indicated that copying had occurred. The frustration of marking copied work increased with the new curriculum change in the province. The increased emphasis on student-generated work, by way of increased laboratory activities, case studies, and greater use of a societal focused curriculum (debates, critical readings, and role-playing scenarios) had increased the amount of teacher marking. Teachers began to question the value of their increased marking. Henry, prior to using portfolios, pointed out:

> It seems that every second night I used to take 30 or more individual laboratory reports home. It took me at least 2 hours to mark them, but I really wondered about the benefit. [pause] Most didn't even listen when we took up the laboratory on the following day. Students would flip to the back (of the laboratory report) to see their mark and then tuck it away in their binder. Lots (referring to the students) wouldn't even listen to the review the next day.

Henry's observations were supported by a number of the teachers involved in the action research project. Henry indicated that students who received a mark of 9 or 10 out of 10 believed that their answers were the only answers. Their motivation to listen to alternate explanations was greatly reduced. When their answer was deemed *"right"*, Henry indicated that the check mark on the assignment tended to end the discussion. Alternate answers were dismissed as "wrong" without proper consideration. Multiple perspectives weren't recognized as possibilities.

Students interviewed at Eastern High School, supported Henry's perceptions. The mark, according to one student, was what she valued most. Once that mark appeared she found little interest in looking at the assignment once again. The finality of the grade was expressed by both high-achieving and low-achieving students. Two students, who had received unsatisfactory grades, indicated that they were often either too embarrassed or discouraged to follow the discussion. Shawn, a student who had repeated science 10 twice, explains:

> In many classes students that get good marks flash them around so everyone can see them. This happens even when the work isn't theirs. It's kind of a ego thing. Because I don't copy work and because I don't always put in a great amount of effort, I usually get one of the lower marks. I sort of hide the write-up in my binder.

Teacher perceptions that few students, prior to using portfolios, reviewed their laboratory write-up once it was marked, were supported by students during interviews. Kevin was repeating the science 10 course and now achieving an acceptable standard.

Interviewer: Did you ever re-read your answers?

Kevin: No because I am usually hiding my work!

Interviewer: What happens if your mark is acceptable.

Kevin: Well it didn't happen all that often. I guess I usually just sit there in relief. Once the marks are recorded why bother - like, it's over!

Interviewer: Has the portfolio changed anything?

Kevin: Yeah, I usually add something from the lab discussion. In some cases I have even gone home and added some more stuff. Except, like, it gets a little too messy now with two or three different colors of pen. I even end up writing in the margins now.

Portfolios and revising answers

The action research group, in moving toward portfolios, no longer evaluated each individual laboratory report. The laboratory write-up was scanned for the completion. The check was performed to encourage student accountability and provide the teacher with feedback on the perceived difficulty of the assignment. A non-evaluative seminar, where students verbally share answers, followed the visual check. The new emphasis had changed from the classroom from one determining answers into a forum for a discussion of results and interpretations. During the discussion, students listened to answers given by their peers and re-evaluated their initial ideas. Students were encouraged to refine their reports and change many of their answers in light of the discussion. The following excerpt was taken from an interview with Henry and Albert after working with portfolios for one semester. Both are experienced teachers from Northern High School. They commented on how the movement away from individual reports changed their classroom atmosphere.

Albert: I find that the *'top'* students really listen to other students, not just for the answers, but for the way in which other students express their answers. [Pause] Some of the less gifted students listen for the answers, but they have to sort out the good answers from the *'not so good'* answers.

Henry: Yea, they really do listen to each other. That did happen before! Some even find that students who don't do well on exams have a great deal to contribute. They may not be good at exams, but they have tremendous insight. You know the kind of kid, the one that just can get it down so we can understand it. [pause] I think some of the class looks at these people differently.

Interviewer: How has this changed your classroom?

Henry: Well the whole notion of cheating has changed. Maybe even more important - they no longer look to me to confirm every answer. I usually say 'well, Amanda, what do you think?', or 'Tom, does that sound right to you?' Only rarely do I interject now. Before, I was always the final word.

Interviewer: Why were you the final word?

Henry: I guess its because my role was to explain why an answer wasn't right. No one ever wanted to know why their answer was marked correct.

Albert: You know it's more than just that. I think in some ways they see me differently.

Interviewer: In what way?

Albert: Well I don't feel like the cop, always trying to catch cheaters. Sharing is wide open. I expect them to add to their answers. It isn't cheating anymore. They still have to decide which answers are worth using.

Henry: I agree and this also increases my expectations. I tell them right off, I will no longer accept a blank answer. If they don't understand the answers thrown about the class, they still have time to come to me for help. (The portfolio may not be due for a number of days).

With the seminar structure, the teacher's focus shifted from evaluator to facilitator of the discussion. Formal assessment occurred later by way of portfolios. Students submitted their refined laboratory reports, at three times throughout the semester. Laboratory reports, for most members of the research group, were accompanied by journal entries that may have included concept maps, responses to literature, or an analysis of difficult exam questions. In addition the portfolio included sample unit or chapter questions that demonstrated the growth in their learning. In all three action research groups, the portfolios comprised work indicated as compulsory by the teacher and some chosen by students. Portfolios were marked in four categories: Development of Scientific Skills, Scientific Accuracy, Communication, and Language Conventions. See appendix 6J and 6L for a sample portfolio and appendix 6K and 6M for a sample scoring rubric, used at Eastern High School. (It is important to note that a number of changes to the scoring rubric evolved over time and that not every teacher used an identical scoring guide. Teachers were encouraged to modify and personalize the portfolio templates.)

Policing copied work and the nature of science

How do teachers make sure that the laboratory write-up actually represents a particular student's work? This question was raised at the initial meetings for all three action research groups. A number of teachers feared that the movement toward portfolios and increased emphasis on group work, by way of performance-based-assessment, would only amplify this frustration. The following questions surfaced during early meetings. Would some students choose to do their portfolios only a few nights before they are due? Are we creating bad work habits? Would laboratory activities that are not identified as compulsory be disregarded? How do we ensure that all of the work is really being done, when students are given choices?

Traditionally, many teachers in the research group used shorter labs that could be written up within the class period. By insisting that students complete the laboratory within the class time, the problem of copying and increased marking time could be reduced. However, this *"cops and robbers"* approach to prevent *"beating the system"* is not without casualties. Laboratory work is often focused upon getting the correct answer. The time to reflect upon the data, researching for more information, and synthesis is severely reduced. Would students in this setting really work as scientists?

Attempts to prevent cheating, if they become the primary focus of a program, greatly distort the nature of science. Science, rather than an open-ended pursuit of questions, is reduced to a set of answers, based on definitions or formulas. Rather than promoting creative thinking, investigations ensure a commonality of answers, where commonplace language serves as an indicator for determining the "correct" answers. Skills are not assessed and discussion is minimized in favor of efficiency. The goal becomes getting everything done before the class ends to prevent cheating, rather than analyzing ideas and seeking alternative answers.

Prior to using portfolios, Gordon explained that students' focus had traditionally been on a single-correct answer, expressed with an economy of words, regardless of the intentions of the teacher. His desire to initiate discussions, that juxtapose alternate conclusions and employ synthesis and evaluation skills, were most often short-circuited by students who sought single answers. Gordon explained, "all answers began to look the same. It was like correcting the same lab 35 different times."

Gunstone (1988) indicates that one of the most troubling failures about science education is that students do not understand that scientific ideas are enhanced through a process of sharing, debating and consensus building. The prospects of isolating students and instituting a time-limiting factor do little to promote this view of science. These activities tend to be highly focused and answer driven. By contrast, the use of portfolios promotes sharing and reflection.

Although a single "correct" answer can be checked quickly, this approach is not without some loss. Cohen and Harper (1991) indicate that some of the best science occurs when the teacher does not know the answer. Science, rather than merely an exercise in problem-solving, should take on the added dimension of problem posing. Hypotheses would be formulated, theories proposed and then tested. If the hypothesis was falsified, the supporting theories would have to be restructured. In turn this should lead students to formulate other hypotheses and begin yet another round of testing.

Teachers in the action research project acknowledged that cheating could never be eliminated; however, a few members of the action research group expressed a concern that portfolio assessment might contribute to the process of copying the work of others. Most disagreed. Henry pointed out that many elements of the portfolio could not be copied. He explained; "How could concept maps be identical?". Albert pointed out that "reading someone else's work is not always bad, just as long as you write down some of your initial thoughts". He commented that "even scientists change their minds when they find a better answer". Every teacher within the action research project devised some strategy to check to make sure that the laboratory report was completed prior to the seminar class. In some classrooms, students who had not completed their initial work were not permitted to remain in class for the seminar. In another class, a commonly-occurring disregard for completing laboratory assignments resulted in a parent conference. Each teacher devised some strategy to ensure that students accepted personal responsibility for completing their work. Once an initial answer was recorded or the teacher received evidence of some thinking prior to the discussion, the purpose of monitoring for copied work had been removed.

During interviews students supported the movement away from teachers assigning next-day laboratory reports and focusing on monitoring student work for evidence of cheating. Portfolios removed the surveillance for copying. In addition, students accepted the idea that each of them should be responsible for completing their work. The following excerpts are taken from different student interviews.

Shawn: Well as you know I didn't do the first portfolio. And it really cost me. I found that I couldn't complete all of the work the day before it was due.

Interviewer: But I saw that most of it was already completed.

Shawn: It was but I hadn't put it all together with the revisions. Even for those people who tried to copy parts of it, it was impossible to complete. You just can't get it all done in one or even two nights. This time, I am a lot more organized.

Leah: I know some people who copied parts from others. But these people are the ones who always leave everything to the last minute. If you leave the portfolios to the last minute, you just can't get it all done.

Natasha: I really don't get bugged by students who copy. They usually don't get their journals done, so I think you know that they haven't done the work. I don't think too many copy things anyway.

Lisa: People will always try to take the short cut. Portfolios don't make it any easier, but it really doesn't stop them. The thing with portfolios is that you can't even copy it all in one night. Those people who don't get their work done aren't likely to take three nights to copy the work from another person. Why even bother copying it? You can get any of the answers in the seminar anyway. It's not like you do it one night and hand it in the next day.

Students abdicate ownership

Before using portfolios, members of the action research group had asked students to hand-in individual laboratory reports the day following the experiment. Thirty-five reports were corrected the following night and a list of correct answers was provided the following day. Should the teacher have multiple classes, it might require an additional day or two before the laboratory activity is evaluated, problems are identified, marks are recorded, and the investigation is debriefed. At each stage the student becomes distanced from his or her work.

Sanger (1990) drew on work conducted by an action research group in England and indicated that students had explained that they believed that their work, once marked by teachers, was no longer their own. The teacher took ownership upon marking. The abdication of ownership, according to Sanger, caused students to disregard teacher comments and grades that were placed on their work as belonging to the teacher. Jane, as a first year teacher, explained how she initially fell into the trap of taking over responsibility for student work.

> I found that I was marking everything. Even explaining answers that were incomplete. The first few labs were a sea of red. Many contained more of my writing than theirs. [Pause] It seemed that the more writing I did, the less likely they were to even look at it.

She went on to explain:

> After doing all of that work and trying to show them the answers, I got really resentful if they didn't look at my comments. I began to feel like I was doing the work and not them. Rather than sitting together as we have talked about, I saw assessment as them and me. It was something I was doing to them and then trying to justify it.

Following one interpretation of a student's transcript, Sanger's (1990) action research group examined their own motivations for marking assignments. Most members of the group found elements of appropriation of student work within their own practice. These included: marking in student books as a means of taking ownership, not offering students a chance to respond to the marking — as a means of keeping control, and by numerous forms of control when freedom of choice or negotiation were supposedly being offered. During an interview Gordon supported Sanger's finding with this comment:

> Sometimes you mark an answer incorrect because it is incomplete. Actually, it isn't wrong, but you're just not sure what they are saying. You end up reading something into it that just isn't there, and you assume that the student is off base. Unfortunately, they don't have an opportunity to respond to your interpretation. That is one of the reasons why I now keep journals. Many students who will not say something will write it. In a few cases it has caused me to look at what they are saying more carefully.

Portfolios and student ownership

The action research groups assigned portfolios that identified a blend of compulsory work and optional work which could be selected from a list of assignments that they had completed. In an interview, Gordon pointed out that students felt they were participating in their own assessment, because they could make some decisions about what is to be marked. Henry indicated that it also helps inform the teacher about what students find most important and challenging. Larry indicated that having students choose what they felt was important likely enabled him to begin asking students for their input in constructing a scoring guide for a performance-based-assessment activity in which science 10 students were to determine which hot plate was most effective. Similarly, Henry explained how having students make some choices in the portfolio may have influenced him in venturing further in asking students in the class to help him develop a scoring rubric for two activities in chemistry 30. Henry pointed out that he had never asked students to become involved in their own assessment in such a manner in over 20 years of teaching.

Only one of the teachers expressed apprehensions about students choosing less challenging assignments. Jane responded to this concern by explaining that the choice of questions also conveys something about the student. She remarked that those students, who search for the easy questions, are telling you that they don't want to be challenged. Jane's conclusions were supported by Anita, a high-achieving science 10 student in a class that Jane did not teach.

Anita: I take a long time to select tough questions. I want you to know that I understand the stuff that is difficult. Some of it, I'm not even sure of and I change my mind. But I feel that is OK, too. I want you to see what I am thinking and how it has changed sometimes.

Interviewer: Do you think that it is possible to get every answer correct in the portfolio and still not get 100%?

Anita: Of course! If you only pick easy things to be marked. This doesn't mean that you really understand science.

Interviewer: But isn't it still right?

Anita: Yes, but you want to show that you can get the tough stuff. That tells the teacher something about my ability. If someone complains just because it is right, then they don't really understand it. I mean, you said that the portfolio should provide evidence of our learning. Exceptional marks must be given for exceptional effort and thinking, not just correct answers.

Interviewer: Would you say choosing what to put into your portfolio has something to do with how it is marked?

Anita: That is it. If you don't even know what is challenging you can't expect a good mark.

Anita's conclusion, that choosing what best reflects your work as a science student carries a responsibility, was echoed by a number of teachers. Gordon spoke about a responsibility for making choices. Marvin indicated that he liked students being responsible for showing him what they understood, rather than him trying to decipher how they responded to our questions. Henry commented on how the opportunity to make a choice placed the onus back on the student.

Henry and Albert commented about changes that they would make after working with portfolios for one semester. Both indicated that they intended to give students more choices the next time they used portfolios. They believed that their need to control stemmed from not knowing what to expect. They explained:

Henry: The next time I do it, I will provide more choices.

Interviewer: Why?

Henry: It gives me a better idea of their creativity. Also, students have really bought into the fact that they can determine what I look at.

Albert: I found the same thing. The first time I was really tentative, so most of the assignments were compulsory. I didn't think they would choose difficult things.

Henry: That's right! I wanted to make sure that all of the hard things were compulsory. Now I want to see who chooses those things. That tells me something about their attitude. It is something that I had no idea about how to evaluate before this. After all how do you assess most of those things in the attitude column (referring to the science curriculum guides).

Henry's comment that many of the attitudinal objectives, identified in curriculum guides, could be understood when students began to choose what they thought was important was also echoed by Gary and Gordon in separate interviews. Although both expressed a frustration with reducing attitudes to a mark, both commented that student choices provided a vehicle for assessing attitudes. Gordon commented that an attitude component was reflected in his assessment strategy, because of the manner in which he was using portfolios.

Holding seminars in which students read segments of their laboratory report or responded to the answers read by other students changed the direction of discussions within the classroom. Gordon and Thad, in separate interviews, commented on how students looked to other students to evaluate their answers. The teacher was not the focal point of all discussions.

Jane also concluded that the students' critique each others answers, during seminars, made a smooth progression toward peer assessment of a murder mystery project by science 10 students. The Murder Mystery Project is now used in all science 10 classes at Eastern High School. In the project students write a mystery using different science clues. See appendix 7B for an outline of the murder mystery. Upon completion of writing the mystery and assembling clues, another laboratory group attempts to solve the mystery. The group that attempts to solve the mystery is then asked to perform a peer assessment of the mystery. See appendix 6A for the peer assessment template for the murder mystery. In addition, each member of the group that wrote the mystery was asked to complete a group assessment for the manner in which they cooperated. See appendix 6B for the assessment template on group cooperation. The following excerpt helps explain how Jane used peer assessment.

Jane: I used three components in determining the group's mark. First I looked at the peer assessment. Next I read over the mystery and jotted down some notes. But I didn't assign a mark. Then, I looked at the way the group assessed themselves for cooperation. From that I determined a mark.

Interviewer: Did you add the three marks together? How was it done?

Jane: Initially, I thought of doing something like that, but it just didn't make any sense. You can't say cooperation is 30% and peer assessment is 30%. I just used all of the information to get a holistic mark.

Interviewer: How was that received?'

Jane: Actually, much better than I had ever anticipated.

Interviewer: Didn't anyone object?

Jane: No. They all seemed to think it was fair. Actually they had lots of information on what was good and what needed improving.

Eisner's (1985) notion of connoisseurship is expressed by Jane. By interpreting the information provided, Jane employed expert judgment during assessment. By coupling student assessment with teacher assessment, she made some advances toward the hierarchical structures of traditional assessment strategies.

Jane's conclusions were supported by four science 10 students during interviews. None of the students expressed any concerns about the fairness of peer assessment, nor did they feel that the final mark did not reflect their work.

Creating an environment in which students participated in and gained some ownership of their assessment helped reduce the bickering over marks, a problem identified by teachers prior to using portfolios. Jane and Larry expressed the idea that providing students with choices diffused some of the problems related to justifying a mark. Their perceptions were supported by Shawn (a third time in science 10) and Amanda (an honors student in science 10), during separate interviews.

Shawn: Well I think portfolios are fair. It is difficult to even argue with an assignment that you choose to present.

Amanda: I know that I like to have choices. You know that I am quiet in class, but just because someone doesn't say something doesn't mean that they agree. This allows me time to really develop things that I am interested in. I can go further and I don't mind putting in the extra work. Sometimes I work really hard and get an 85% or I just put in a little work and get an 80%. You know the mark doesn't always show the work. In some courses, like Social, we seem to be turning in an assignment or getting a quiz every second day. I really get sick of being evaluated.

Amanda's point about being evaluation weary was expressed by other students during interviews. Leah, a conscientious hard-working grade 12 student, expressed frustration with constantly being measured. She indicated that after a while you get so many quizzes and hand in so many assignments that they lose meaning. She explained:

> After a while you don't even worry because you just say to yourself ' I will make it up next time' or something like that.

Tobin, reflecting upon a case study in which a teacher used assessment as a dominant means of motivation, states:

> A close relationship between assessments and academic work did not produce an environment in which students were encouraged to take risks since almost everything written by students was graded. Students in all classes had few opportunities to practice skills in a formal sense without the threat of a grade (Tobin, et al., 1988) p. 439).

Orlich (1980) warns about the dilemma of using assignments to increase student accountability. The coercion of a mark should not be used to provide the impetus for ensuring that students complete their work. The value of doing the assignment resides outside of the learning accomplished by the student. The value is often fixed with the grade given by the teacher. Not only does the student relinquish the work but also the valuing of his or her own work. To make matters worse, the teacher also loses under this system. The teacher accepts responsibility for grading the assignment, identifying where the student has gone wrong, and providing a guide or key to the correct answer.

The inevitable spiral of assigning work, grading the work, and the administrative tasks of recording and charting student performance, leaves teachers with little time to spend on preparation and, hence, their lessons become even less interesting or less relevant. To ensure that students comply with their objectives, even more work must be assigned and the cycle continues. Progressively, student assessment consumes even more time and the role of the teacher is drawn even further from the student.

Hidden consequences of accountability

Pressures to improve test scores in the absence of serious, parallel support for instructional improvement is likely to produce serious distortions between what is valued and what is evaluated (Herman, 1992a; Wolf, Bixby, Glenn, & Gardner, 1991; Zielder, 1991). Superficial changes in instruction to improve test scores are not likely to result in meaningful learning (Mehrens, 1992; Simmons, 1992; Wiggins, 1989). As a result test scores do not represent broader student achievement, but only content formats, included on the tests. A more positive picture is achieved when assessment strategies are focused upon providing models of authentic skills (Shavelson, 1992; Wiggins, 1992b). Some evidence suggests increased student performance can be associated with using more authentic assessment strategies, although a change in assessment practices is but one of several factors that affect student and teacher performance (Herman & Golan, 1992). Authentic assessment attempts to link curriculum goals with assessment by broadening assessment strategies.

Important Points :

- Poor thinkers and problem-solvers differ from those classified as good thinkers and problem-solvers not so much in their skills but in their ability to apply them to the appropriate task.
- The social context in which problems are introduced also affects the approach to assessment. Real-life problems require that people work together as a group. Independent problem-solving is de-emphasized in favor of collaborative and cooperative ventures.
- Assessment criteria must be expanded beyond validity and reliability to evaluate the quality of assessment strategies. Consequences, fairness, transfer and generalizability, cognitive complexity, content quality, content coverage, meaningfulness, and cost/efficiency must be considered to determine quality (Herman, 1992a).

Curriculum fit and evaluation are taken-for-granted assumptions when individuals adopt a technical approach to research. Clearly, an approach for viewing assessment, other than the technical (end-means) approach, is needed. The motivation for looking at assessment can not focus on increasing test scores. Determining assessment quality requires that teachers, administrators, and school boards work toward common goals. These goals should consider how assessment can support learning and curriculum intent rather than delimit it.

What happens when students score in the 80s on exams but only get 60s on other types of assessment? Teachers in the action research group expressed an early concern that the introduction of new indicators would disrupt the 'pecking order' in their classrooms. Students who usually score above 80% on written tests have the least to gain and the most to lose. Not surprisingly, I found that, in many classrooms, those are the very students who have most influence with teachers. Objections about new approaches, from students who do well on exams, are often heard louder. In many situations these motivated students also received the greatest support from parents. Teachers expressed a concern over these parents expressing displeasure with the new approach.

A linked concern was that the new indicators would not be supported by diploma marks. Teachers were most reluctant to try alternative assessment strategies in diploma courses. They believed teacher-assigned student marks would be compared with diploma marks as an arbiter of how effective their assessment had been. Science 10 became the most popular trial balloon, with the 20 level subject courses following closely. Even after a number of successes at the 20 level, many of the teachers were reluctant to change the manner in which 30 level subjects were assessed. The adjustment can not be rushed. Teachers need time to develop confidence in new assessment strategies.

Teachers' difficulties in abandoning old paradigms

Trying new assessment strategies requires abandoning tried and true techniques. Marking individual laboratory reports, only to have students compile them at a later date for portfolios not only dramatically increases a teacher's work load, but also defeats the very purpose of using portfolios. However, abandoning a repertoire of refined methodologies is especially difficult for experienced teachers. These techniques have evolved along with the teachers and in some sense become part of the way in which they view teaching. Thad explained his reluctance in giving up some of the assessment techniques that have become part of his repertoire. Thad, explains:

> It is easier for newer teachers, I think, because they haven't worked with an approach and tried to tinker with it year after year. You make a little change here and another there. Eventually, you become comfortable with it. You begin to trust it!

Later in the interview he went on to explain:

> The approach becomes yours. You would hate to think that something that you have used for 25 years is wrong. (There is always some new and improved teaching approach that someone is trying to sell). If it no longer works with a larger number of students, you just conclude that the students lack motivation. They aren't as mature, less academic – you know.

Thad points out that he would likely reject any pre-packaged approach, as a method. He has seen too many come and go. Any changes in assessment require the abandoning of some of our traditions without the rejection of everything that we have developed over the years. Assessment strategies that are to be abandoned are not deserted because they are no longer "correct", but because they are too limiting.

Not everyone needs to start in the same place and at the same time. Thad spent one year observing student reactions, and talking to other teachers before attempting to change his assessment approaches. He indicated that the opportunity to observe others and acclimate was essential. Henry, another experienced teacher, indicated:

> I told the students that I was going to try it for a semester. I didn't think they would like it. I tried it only for the second reporting period. I felt that I had to get to know them first. . . . Later we decided to do it for the third term.

Successful change means that everyone who wishes to change assessment strategies should set their own timetable. Those interested in bringing about change must be sensitive to the fact that classrooms have different ecologies. It is not a race to some pre-determined goal. Most importantly, teachers must be allowed to select opportunities in which they believe they can successfully implement change.

Teacher concerns about changing assessment strategies

Teachers willing to venture into a new way of assessment must be prepared to deal with uncertainty. The dilemma is underscored during my interview with the teacher identified as Marvin. In describing why he had concerns about using portfolios and performance-based assessment Marvin indicated the following:

> I don't like not knowing what I am going to get. How can I be sure that they completed the assignment themselves? With exams I know what they understand. I can even usually predict the mark that the students will get, at least within a range. It is much fairer. All I have to do is add up the marks. There is no room for mistakes, as long as I add up the marks correctly.

Marvin went on to explain that sometimes, strong academic students didn't do well on performance-based assessment. Some students with good listening and reading skills do well on exams, but may do poorly on portfolios if they lack organizational skills or motivation. Yet, he explained, they know the material. Marvin explains:

> Exams and not portfolios or performance-based assessment provide better indicators of how students will do in future courses. [later he explained] Universities use exams. If you can't write the exams you won't be able to continue. Like it or not we are preparing them for the next level.

One of the strongest concerns focused on the use of portfolios. Unlike individual labs, portfolio assessment does not provide constant feedback. Mary expressed an early concern:

> Yes I checked their homework, but they don't really know how they are doing after every lab write-up. If they had a mark of 5 out of 10, they would have an idea. Merely discussing the things to look for isn't enough.

Mary indicated that portfolios only delay assessment. What happens when the first feedback on marks only comes after 10 or 15 activities are completed. She expressed concerns about whether it was fair to students.

Another concern, linked specifically with portfolios, was that, by providing choices, not every student would have the same thing graded. Marvin and Mary expressed concerns about whether or not a representative marking system would be fair. What happens when two different things are marked? They expressed concerns about accountability. Would students perceive their assessment to be fair?

Student perceptions about teacher concerns

Student perceptions were gathered from a survey of 150 students from 6 different classes taught by three different teachers and 16 students interviewed in groups of two on three separate occasions. Student responses were collected not only to trace their reactions to our work on alternative assessment strategies but also to validate teacher perceptions about student needs.

Teachers' concerns:

1. Students need immediate feedback about their labs. Portfolios delay feedback.

The students agreed with this statement; however, in interviews a number expressed that they did not need a mark to tell them how well they were doing. Anita summed up this query by saying:

> Well I know if I understand it. The seminar for one helps me know if I get it. The mark just tells me more about how the teacher is rating me. It's really not the same thing.

Teachers at Southern school abandoned the use of portfolios, after one semester, largely because they felt it delayed assessment. Teachers at Northern and Eastern employed other strategies to provide students with ongoing feedback. Henry and Albert coupled portfolios with performance-based-assessment to keep students informed on an ongoing basis. Gordon, Larry, Marvin, Thad, and Jane all provided non-graded feedback on the first two laboratory reports. Every teacher involved in the action research group at Eastern and Northern High Schools also provided students with a sample write-up from their first laboratory activity. The sample write-up was designed to serve as an exemplar. Gordon indicated that the next time he used portfolios he wanted to employ more quizzes to provide students with a better record of their progress. He felt this was especially important for grade 10 students, unfamiliar with the way portfolios were used in science classes.

No clear indicator for student preference for portfolios over individual laboratory reports was identified in the survey. Students do like to know where they stand at all times in the course. Despite complaining about the wear of constant assessment, it also provides them with security. Students, like teachers, do not like surprises.

Henry indicated that he found that many of the conscientious students preferred portfolios because they were rewarded for things that they had always done. But also cautioned that many of those, especially gifted at writing exams, found it too time consuming. They would have preferred to have their entire mark determined by way of formal exams.

Survey numbers are recorded in % of the population (n= 150).

Descriptor	Agree	Agree most of the time	Neutral	Disagree most of the time	Disagree
1. Portfolio assessment is worthwhile.	52.7	22.5	16.2	1.4	7.1
3. I prefer individual lab write ups that provide immediate feedback.	16.5	24.7	34.5	15.5	8.8
4. I prefer a seminar after the laboratory activity or case study where we have an opportunity to listen to the answers given by other students.	35.9	27.4	21.8	5.6	9.2

Teachers' concern:

2. Exams provide better indicators of student understanding and better predictors of future success.

Students did not support the view. In interviews many were able to explain why certain students have great difficulty writing exams. Kevin explains:

> I have never been good at exams. Before the exam I feel my throat get tight and I have trouble swallowing. In grade 8 I just waited to fail the test. The portfolio is altogether different. I have time to ask you to explain what the question is asking. I can read a little, talk to some else and then write down an answer. You just can't do that on exams. They are just so much pressure.

Leah echoes Kevin's concern:

> I once heard that Einstein was not all that good at exams. I learn by talking to others, not just by sitting alone. Exams also have so much pressure that I just forget things that I know.

The survey results show that students recognize that portfolios and exams use different indicators.

Survey numbers are recorded in % of the population (n= 150).

Descriptor	Agree	Agree most of the time	Neutral	Disagree most of the time	Disagree
13. I believe that the marks for the portfolio reflect those of the exams. Students who get high exam marks do well on portfolios while those who do poorly on exams always do poorly on portfolios.	9.2	15.5	27.5	23.9	22.5
14. Individuals who always do their work, but have trouble writing exams, benefit by portfolios.	53.5	29.6	12.7	2.1	1.4

18. I believe that exams provide a better indication of what students have learned in the course than do portfolios.	14.1	18.3	31.7	16.2	19.0

Teachers' concern:

3. Alternative assessment strategies will inflate marks.

Interestingly, most students believed that their portfolios would help their class mark a lot more than it actually did. The teachers' and students' perception that portfolios and performance-based-assessment would raise marks substantially was only moderately supported by doing a mark analysis for six different classes. Marks on portfolio assessment was on average 3.2% higher than exam marks, while performance-based-assessment marks were on average 5.3% higher than exam marks. As a point of comparison, marks on quizzes were on average 4.8% higher than the marks on larger exams.

Descriptor	Agree	Agree most of the time	Neutral	Dis-agree most of the time	Dis-agree
11. I believe the marking scheme used for the portfolio is fair.	37.3	36.6	14.1	7.0	2.1
12. I believe that the portfolio mark will lower my course mark.	14.1	10.6	26.1	19.1	30.0

It is important to note that students' general acceptance of portfolios did not occur because they foresaw it as less work. During the interviews a number of students explained that they took an unusual amount of time deciding which assignments best represented the best work that they had done.

Descriptor	Agree	Agree most of the time	Neutral	Dis-agree most of the time	Dis-agree
5. Portfolios require less work than individual laboratory write ups and hand in questions.	4.9	8.5	19.0	28.9	37.2
9. I find it time consuming to select work for the portfolio.	14.1	25.4	38.7	9.9	10.6

Teachers' concern:

4. Because students have choices, different things will be used to determine a grade. Everyone should have the same assignments marked.

Students did not support this idea in either the survey or interview. Lisa, a high achieving grade 12 student, explained:

> We are all supposed to have different strengths and learn in different ways. Why are we evaluated in the same ways? In some ways it is like deciding who you want to succeed in advance. [pause] They are the people who are good at writing exams.

An interesting anomaly arises when analyzing the survey. (See questions #2 and #6.) Although many students believe that they should have choices in which assignments are to be marked, some of these same students also indicate that fairness in assessment is achieved when everyone does the same things. Students, like teachers, live in a rational world, where consistency and repeatability are desired goals. Unfortunately, the incompatibility of recognizing individual strengths and the movement toward conformity or standardization is not often recognized. Many students assume teacher expect teachers to achieve both goals simultaneously.

Descriptor		Agree	Agree most of the time	Neutral	Dis-agree most of the time	Dis-agree
2.	I believe that students should have some choices in determining which assignments should be marked.	40.9	23.9	23.9	5.6	4.2
6.	Everyone should hand in the same material to be marked in order to get a fair assessment.	29.6	22.5	18.3	13.4	15.5
8.	We shouldn't use portfolios because it only provides a sample of our work. All of our work should be marked.	3.5	6.3	19.0	35.9	33.1
19.	The teacher should choose all elements that go into the portfolio.	9.9	11.3	29.6	20.4	27.5

Benefits to students

1. The portfolio as a source of review.

During interviews, students identified portfolios as an excellent source of review. Interestingly, this potential benefit, identified by nearly every student, was not mentioned by a single teacher during the planning stages of the research. This idea was later confirmed in the survey:

Descriptor		Agree	Agree most of the time	Neutral	Dis-agree most of the time	Dis-agree
7.	Portfolios serve as a good source of review.	43.7	28.9	13.4	7.0	5.6
10.	I tend to read over labs, case studies, questions, and journal entries prior to handing in the portfolio.	36.6	33.1	14.1	6.3	8.5

2. Journal entries provide students with an opportunity to engage in open-ended hypothesis formation.

Natasha made the following comment during an interview:

> I went back to my initial answer four different times. It seemed like every time I learned something else, I changed my mind. [later] It looks kind of odd now because there are notes in the margin.

Most students, once they have adjusted to open-ended questions, do not worry about the fact the a single "correct" answer may not be given, or that the teacher may not have an answer. In the survey, students rejected the idea that journal entries require pre-defined answers in order to be useful.

Descriptor	Agree	Agree most of the time	Neutral	Dis-agree most of the time	Dis-agree
17. Journal entries have no place in the portfolios because they have no single right or wrong answer.	7.0	13.4	21.1	29.6	26.8

Summary

New strategies for assessment provide exciting opportunities, not hard-and-fast solutions to old problems. The action research project was much more than collective problem-solving. Documenting the successes and failures of the school-based action research groups, as we worked to implement assessment strategies that better reflect learning, have provided important signposts to others interested in similar problems. Perhaps the most significant part of the study is the focus on student perceptions of assessment strategies that were designed to provide more democratic classrooms. The list below outlines some of my journal entries in overcoming hurdles to implementing new assessment strategies

- Expect to work with a fair amount of doubt. The removal of commonly held reference points makes success and failure more difficult to evaluate. The frustration of not knowing was greatly alleviated by working in a group with other teachers. Sharing experiences provided direction and moral support. The school in which sharing was limited reported less success with bringing about changes in the manner in which they assessed students.
- Do not expect every student to embrace new approaches to assessment. The question about who is rejecting the new approach is most important. Is it the majority of students, or only the most vocal? Not every student will benefit by changing assessment

approaches. Expect the changes to disrupt the hidden stratification system that exists in your classroom.

- A preset plan or method can not be given to teachers to follow. Teachers will reject the technician's role, especially if it challenges past practices that have become rooted in their personal beliefs. Although sample templates for performance-based assessment and portfolio design were shared, teachers exercised the prerogative to modify, restructure or completely reject the strategies devised by colleagues.
- Teachers must set their own timetable for change. Forced change, according to a number of teachers, will be met with resistance.
- Bringing about change is most effective when students are asked about the type of changes that are being made. I believe that student interviews and formal or informal surveys by the teachers involved in the action research project greatly affected the outcome. A number of students expressed their appreciation and sometimes surprise at being asked for their opinions.
- Parent, student, and administrative support can be very helpful; however, it is important to note that none of the action research groups reported any interference. The much anticipated barrier never manifested itself in any of the school sites.
- Students are not sold on teachers only marking representative samples of their work.

Chapter 6: Conclusions and Implications

Overview

A quote, reportedly found on the wall of Albert Einstein's study, encapsulates current assessment practices in most schools.

> Not everything that counts can be counted and not everything that can be counted counts. (Ryan, 1994, p. 1)

Aspects of curriculum that easily lend themselves to quantification, are tested most often. Subsequently, what is tested defines what is valued. This allows assessment to define curriculum and ascribes "winners" and "losers" in classrooms. It validates things that are said to "work" and it detects other things as "failures".

In the conclusion I will return to the research questions by linking the philosophical underpinnings of the thesis with my interpretations of the action research program. I will also examine the many unanswered questions that the research has brought to light and provide suggestions for further studies.

Reflecting upon the research questions

What follows is a reflection on the original research questions. The reflections are based on summary notes made during the project and insights gained writing the conclusion. In many ways, writing the thesis has forced me to slow down and re-think many of my interpretations. As an active participant all too often my understanding of the issues have been compromised by devising strategies that met with student or colleague approval. The focus and time was too often spent determining what worked within the classroom rather than analyzing why it worked or what it eliminated by working.

Discordance of intents

1. Are the evaluation strategies being used by high school science teachers to assess students' knowledge of science compatible with the intents of the new curriculum?

Teachers in the action research groups were unanimous in stating that a conflict existed between curriculum intents and current assessment strategies. They expressed great frustration in meeting the curriculum initiatives and reconciling them with an assessment agenda rooted in accountability.

Although members of the action research group praised the attempt to present science within a social context, no one believed that the same message had been conveyed by current assessment practices. During an interview, Gary expressed his concerns by stating:

> The real curriculum is what is tested. To get students ready for these kind of tests, you make them practice. You know, even in grade 10 and 11. What really counts is not their understanding of STS, but how many questions they get right.

Gary's practice of aligning teaching and assessment to meet the demands of standardized testing is supported by research done by Herman and Golan (1992). They documented how standardized testing affects teachers' planning. They reported that scope and sequence is adjusted on the basis of testing. Teachers in this study (Herman & Golan, 1992) devoted considerable time for practice exams and reviews. A similar observation was made by the action research group, especially at the grade 12 level, where many laboratory activities were omitted in favor of exam preparation and practice. In addition, Herman and Golan (1992) noted that many teachers looked at prior tests to ensure a good match between curriculum intents and their teaching. This practice was also prevalent in the action research group. Teachers within the action research even commented favorably about exam banks prepared by Alberta Education's Distance Learning and Curriculum Branches. Teachers indicated that the test questions helped them interpret what they believed was important in the curriculum. No one in the study ever suggested that the questions might contain what was easily testable on pencil and paper tests, rather than what was important.

The power of standardized testing in determining curriculum was identified by Marvin during an interview, early in the research project. He stated:

> In diploma subjects you know what is important. You just have to look at the tests. But it's harder in science 10 and now, even in the new biology 20. Reading the curriculum document or even looking at the texts doesn't always indicate what to do. [Later in the interview] I am not sure about using debates. Students like talking but is it real science. Sometimes I believe it serves little purpose. We test them for facts not their opinions.

Marvin's understanding of assessment was supported by a rationalist framework. He identified things that were important because they were tested, rather than asking if all important things could be assessed by traditional approaches. Debates were questioned because he saw no way of evaluating them. Early in the study, Marvin explained that science was a collection of facts that could be tested. By the end of the research study he accepted debates; however, he never abandoned a reasoning based on a rationalist orientation. During an action research meeting, he explained that debates were excellent ways of getting students to learn the facts. He now saw debates as a motivational technique and as a technique for exposing possible misconceptions that a student may harbor.

Teacher frustration over what they believe is conveyed, by curriculum intents, and what is really valued, by way of standardized testing, is well supported by the literature. Dorans (1994) interpreted a study done by Martens in 1992 and explains that curriculum developers, superintendents, and parents give teachers conflicting messages about what is valued and what should be assessed. Martens noted that the assumptions made by different groups, instead of being shared were in direct conflict much of the time. Most importantly, because the divergent groups did recognize that their own agenda was separate and often in opposition to that of other groups, potential areas of conflict were not identified. The curriculum plan and supporting assessment strategies were inappropriately seen as meeting everyone's needs simultaneously. In her study of the implementation of a curriculum initiative directed at improving student understanding about the nature of science, Martens found that science directors and teachers believed that student understanding about the nature of science could be accompanied by lower scores on a standardized test, if the exams continue to focus on science content as a primary source of knowledge. Taking more time to ensure that students gain a more authentic view of the scientific endeavor reduces the time spent learning factual information. Not surprisingly, the linking of the new initiative with lower test scores brought tremendous pressure on the program. Many teachers abandoned the new approach, despite believing that it improved their science program. The fact that the standardized test did not evaluate many of the aspects of new curriculum seemed not to matter. Classroom teachers were not in a position to challenge the test.

Teachers within the action research group expressed many of the same feelings of powerlessness. Those who challenged aspects of standardized testing feared that they might be viewed with suspicion. Were they attempting to deflect responsibility for lower achievement by students? When a member of the action research group complained about one aspect of standardized testing, the statement of criticism was often preceded by telling the group about the success their students experienced on standardized tests. Many members of the action research group felt compelled to state that dissatisfaction did not arise from personal motives. The idea that a complaint would be tainted unless we could clearly demonstrate impartiality worked as a constant reminder to the facts that we have developed within a world of rationalism. In a rational world objectivity is prized. Values or beliefs, linked to the subject, are met with suspicion, because they suggest bias.

The frustration of meeting divergent and often conflicting needs was expressed by Thad during one of the interviews. He expressed a perceived expectation of the educational community that everyone should be above the provincial average on diploma exams. Parents, he indicated, will accept the notion that university is not for every student, just as long as their child makes it to university. He also explained that there is a movement to make science courses more academic, so that Canadian students are able to compete in a world marketplace. During an action research meeting at Eastern High School, teachers supported Thad's statement. Later they identified concomitant expectations that more students take science, enjoy our classes more, that the drop-out rate decreases, and that achievement continues to increase. They discussed how these issues would not be addressed by changing the forms of assessment.

Support for changing assessment strategies

2. Have assessment approaches changed?

During the action research project, assessment strategies did change. However, the amount of change varied tremendously among the different research groups and between members of a single research group. No common timetable for change should be expected. I found that some teachers appeared to embrace new approaches faster than others. A similar interpretation can be identified in the work of Lantz (1987), who found different qualities among chemistry teachers capable of supporting changing paradigms. In another study, Duffee and Aikenhead (1992) described how personal knowledge supported changes in classroom approaches to provide an STS focus for curriculum. Both studies indicate that previous experiences, divergent views on classroom structure, and the climate of the school community affect a teacher's willingness to institute changes within his or her classroom.

In my study, no single factor stood out as a single indicator to explain why some people were more reluctant to initiate change. Marvin and Thad, both experienced teachers, were slowest to begin changing assessment strategies; however, age should not be viewed as the salient agent in determining a willingness for change. Henry and Albert, two teachers of nearly equal experience, were most willing to try new assessment strategies. It should also be pointed out that a cautious beginning does not mean a slow plodding pace. Some of the members of the action research groups at both Northern and Eastern, who were quick to try new things, made a great number of changes during the first few months and then maintained a "status quo" for the remainder of the project. By contrast, Thad moved very slowly for nearly three semesters and then began using portfolios and re-designing laboratory activities to include performance-based-assessment components. It was my observation that the rate of change fluctuates greatly throughout the school year. Most changes occur at the beginning of a semester when teachers have the most energy.

Originally, I thought that teachers who felt most secure might be the first to initiate changes in assessment strategies. Various action research studies (Allan, 1990; Baird, 1986; Casonova, 1989; Elliot, 1991) indicate that a feeling of security and support are essential for risk-taking. Although I believe that group support is essential, I found many conflicting indicators for security within the study. Three of the teachers at Eastern (Jane, Gordon, and Larry) did not have continuous contracts with their school board. Each had experienced turmoil during the year, not knowing if reductions in provincial funding would permit them to retain their jobs. They were either first year teachers in the school (Gordon) or first year teachers (Jane and Larry) at the beginning of the study. Yet, in spite of a lack of job security, these teachers were extremely open to change. Conversely, some of the more experienced teachers, who had greater job security, expressed concerns that might be linked with security.

Mary, the department head at Southern High School, knew that she had reasonable job security, yet she often spoke of parents and administrators as limiting change. Despite being a long-serving, well-respected teacher within the community, Mary expressed reservations about trying assessment strategies that might provide different indicators for success than those provided by standardized assessment. These same reservations were raised by various teachers at other school sites. I found no strong relationship between the number of years a teacher served within a school and the willingness to institute changes in assessment strategies. Teacher security, in these projects, was more closely related to the support received from the action research group rather than acceptance and recognition by the immediate community. Teachers like Marvin and Thad, who operated as observers from the periphery of the group were more reluctant to change. As Thad began to institute a greater number of changes, he began to interact with other group members to a greater degree.

Determining what is valued

3. Can an action-research group develop assessment strategies that complement STS teaching strategies? What components can be evaluated? What message will we send to students if we don't evaluate a certain component of STS? Will the students value components of the course not evaluated?

Teachers indicated that more authentic assessment strategies were required to support an STS orientation to the teaching of science; however, a divergence of opinions existed about whether this new orientation was really valued. The action research groups at Eastern and Northern High Schools embraced the opportunity to initiate change. A school developed model for change worked within these two schools. A different school ecology at Southern High School and the practice of field-testing government programs limited the self-directed changes attempted by the school. Members of the action research group borrowed and modified existing approaches developed at Eastern and Northern High Schools rather than develop new ones. I believe that the early departure from action research by members of Southern High School can be linked with ownership. Action research, to bring about change, should identify a need and structure a plan.

As Elliot (1991) warns, action research should not entail the surveillance of teachers who apply an externally derived plan.

During interviews, teachers at all three school sites and students at Eastern High School explained that should an activity not be assessed, it would likely be interpreted as having less value. However, teachers explained that debates and some group work was difficult to assess. Interviews with both students and teachers pointed to a conclusion that things that created ambiguity, such as group work, tended to be valued less in deriving report card marks. Who really did the work and should everyone receive the same mark became contentious issues. Students also described concerns over ways in which group work had been assessed in the past. Students described a frustration in having the final product representing the entire mark. During one of the interviews students explained that co-operation should also be valued, if this was one of the objectives of the activity.

During interviews teachers and students agreed that "what is valued" differs from classroom to classroom. Neither group expressed any concerns over the multiplicity of assessment devices or the diversity of standards in different classrooms; however, concerns were expressed about using different assessment devices for different students within the same classroom. Although both students and teachers acknowledged that learning is idiosyncratic, neither group was able to reconcile that different people might be assessed in different ways. A rationalist perspective of fairness supports the opinion that everyone must be treated in the same way and do the same things. The idea that objectivity is inextricably linked to justice forged an opinion that dealing with the subjects, as individuals, would create uncertainty. Following the rationalist tradition, uncertainty was linked with bias, an undesirable quality.

Using teacher practical knowledge

4. What are the teachers' reasons, within the framework of teacher practical knowledge (Connelly, 1985; Duffee & Aikenhead, 1992) or personal knowledge (Polanyi, 1962), for selecting the assessment practices that they use to evaluate student knowledge?

Teachers in the action research groups used practical knowledge to identify common inadequacies of current assessment practices. During meetings at all three schools teachers voiced a concern that current assessment practices disadvantaged certain students. At an action research meeting at Northern School, Albert mentioned the name of a student who experienced difficulty demonstrating the knowledge that he had gained in biology. Upon providing a context for his statement other teachers within the school, who had taught the student, quickly added to the conversation. They indicated why they believed this student experienced difficulties. As the conversation progressed, other student names were offered as a means of contextualizing many reasons why exam-based assessment was unable to demonstrate the knowledge acquired by all students.

In all three action research groups, teacher practical knowledge guided discussions about student needs and assessment. In some ways the teachers' experience at Northern school was unique because it began as a formalized university course. Readings about learning styles and assessment techniques (Alberta Education, 1991; Alberta Education, 1992; Mielnichuk & Peat, 1990) were used in coursework. However, it should be noted that once the research project began, no formalized discussion evolved around readings. Only Henry had ever asked for additional literature, but his request focused on action research rather than more information on assessment.

Research sites at Eastern and Southern school did not build on a common base of literature. Articles providing directions for changing assessment strategies (Erickson et al, 1993; Rodgers, 1992; Madaus & Kellaghan, 1989; Baird, 1986; McLean, 1985) were provided to all action research groups, but as the participants freely admitted, were rarely read. The teachers appeared much more willing to discuss issues than read about them in isolation. No teachers interviewed, with the exception of Henry, indicated that they had done any reading on assessment on their own.

During meetings, many of the teachers interviewed seemed quick to dismiss assessment programs developed by universities or government agencies, as not reflecting the true climate within schools. The most common complaint, that surfaced about the largely unread articles, was that they would not be sensitive to time demands placed upon teachers. Mary, Marvin, Albert, Gordon, Thad, Henry and Marvin all commented about how so many externally derived programs do not acknowledge an increase in the teacher's workload. Once again, teacher experiential knowledge was summoned to support these conclusions. Past school board or government initiatives were cited most often in general terms. I found that a strong consensus of opinion exists supporting the teachers' belief that individuals or groups removed from the context of a school often fail to understand the complexities of change.

During an action research meeting at Eastern High School teachers poked fun at simplistic solutions to complex problems. They laughingly used sarcasm in commenting on how the newest methodology to solve all problems had everything you would ever need, if only classrooms had a certain type of student. It should be noted that the sarcasm was not directed at any group in particular, but expressed their feelings of disregard for those who they perceived had naively structured change for schools in isolation.

Factors that limit change

5. What factors will limit changing assessment strategies?

The teachers who adopted new assessment strategies spoke of developing a comfort zone. For the rare individuals like Henry and Jane, a short discussion about an approach to assessment was more than enough to get them thinking about how they might adapt ideas for their own classrooms. Others looked to informal leaders for support. They wanted to know that their colleagues had experienced some degree of success with these new strategies before they tried them.

Some of the teachers wanted assurances that broadening assessment, to include criteria not formally evaluated, would not create conflict. At Eastern High School, Larry relied heavily on the experiences of Jane before implementing new ideas. Although, Larry displayed more confidence in social settings, Jane clearly became the informal leader among the two first-year teachers. As the action research group progressed, Larry became the implementor of plans constructed by others. Although he was willing to select from a number of different approaches from different resource people, he retreated from the responsibilities of originating his own ideas.

At Northern High School Albert looked to Henry before attempting new assessment strategies. However, unlike the relationship between Larry and Jane, Albert never procured a pre-made plan and applied it, unaltered. Albert always modified and personalized new ideas. In part this may have been due to the fact that Albert and Henry taught in different subject areas.

In all schools the higher the perceived stakes of accountability, the more reluctant teachers were to initiate change. Gary and Mary indicated that they would only use portfolios assessment at the grade 12 level if portfolios proved successful at the grade 11 level. Both teachers spoke of a fear that an untried strategy might unduly affect diploma marks. Interestingly, they spoke of their fear that student marks would become inflated. Hence, they identified grade 12 subjects that have a diploma exam as having higher stakes.

Many teachers at Northern and Eastern High schools independently arrived at the same conclusions and used either the grade 10 or 11 science courses as a testing ground. Only a few teachers, such as Henry, Albert and Gordon began using portfolios simultaneously in grade 11 and 12 subjects. Teachers, at all three schools, indicated that they felt greatest pressure at the grade 12 level because standardized test scores were closely monitored. This observation is supported by the research of Herman and Golan (1992), who reported that standardized testing created an environment of surveillance of teachers and students in those subjects being tested. Doran (1994) draws similar conclusions from a case study completed by Martens in 1992 and indicates that school administrators spend considerable time discussing with teachers, ways to improve test scores from standardized tests.

The message to teachers is quite clear. Accountability is not understood by matching curriculum guidelines or learner outcomes with teaching strategy or the experiences and attitudes of the learner. Accountability is measured. Most importantly, it is measured by those things which standardized test can measure most easily or efficiently. Teacher practical knowledge is not recognized as a worthwhile arbiter in student assessment. Accepting the notion of teachers as connoisseurs of learners requires getting to know and understand how the teacher comes to understand learners, and this takes time. Reading class averages permits a quick survey of many different classrooms and provides parents, administrators, and the news media means for comparisons.

The greatest impediment to widening assessment strategies, to better reflect curriculum guidelines and better present a portrait of student learning, is that teachers do not always believe that these broader assessment strategies will be valued.

Problems created by assessing a broader curriculum

6. What problems will be created as we attempt to move toward assessing technological thinking and social issue components of a science program?

Albert, Henry, Gordon, and Jane identified a shift in classroom structure that places students into roughly defined groups or categories according to ability. Henry pointed out that he believes that portfolio assessment rewards those students who do their day to day work. He went on to explain that not everyone of these students necessarily does well on exams. Henry expressed the idea that using portfolio assessment was valuable, because they disrupted the *"status quo"* in the classroom. However, he did provide a caution. During an interview Henry explained:

> Students kind of have this idea of where they rank in the class. You know, the I am smarter than so-and-so attitude. When you change the way that you get marks, you disrupt this system. This is especially true of grade 12s. It isn't always received well, especially by those who placed themselves at the top.

Interviewer: Why mostly grade 12s?

Henry: Well I think its because they have been together longer. They are also more conscious of marks. Once they have established a pecking order or whatever you want to call it, they don't want it to be disturbed.

Henry felt that it was beneficial for students to begin to recognize talents in others within a classroom. During an action research meeting with student teachers, Henry explained how static assessment devices tended to establish stratification within classrooms. He explained how a 90% student is often more valued than one who gets 60%. The student teachers took considerable time relating personal stories to reinforce Henry's observations.

Henry's warning for those embarking upon new strategies to expect disruptions in the classroom was noted by others at Eastern and Southern High schools. As strengths, not formally recognized, became valued with assessment strategies, some high achieving students expressed displeasure. The notion of grade 12s being more resistant to change was anticipated by teachers at Southern. Mary explained:

> I don't want to try it (portfolios) at the grade 12 level. Could you imagine what would happen if my class average was way out of whack.

Mary and Gary decided not to change any assessment strategy for grade 12 students until they had gained some feeling of comfort with them. A similar finding occurred at Eastern High School. Thad and Marvin were most apprehensive about changing any assessment strategies at the grade 12 level. Although portfolios were eventually used, it occurred well after they had garnered some confidence in their worth at the grade 11 level.

Teachers at all three research sites predicted some challenges from parents and administrators. The much anticipated confrontations never materialized. Whether teachers intuitively delimited change before parents and administrators restricted classroom autonomy or whether they had misunderstood the beliefs of parents and administrators was never addressed. The principals at Eastern and Southern high schools indicated that they had no concerns with changes to assessment strategies both before and after the project ended. The principal at Northern was not asked about concerns following the project; however, Henry the department head at Northern High School, indicated that questions about any adverse impact of the action research study did not surface during nor one year after the study. To my knowledge or that of the department heads at all three schools, parent perturbation was not voiced.

As reported earlier, the greatest limiting factor was a teacher's ability to deal with ambiguity. The current model for assessment provides certainty because of the heavy reliance on testing. This model achieves certainty by measuring repeatability. Student performance within a traditional testing model is more easily predicted. Students, who score a 70% on one exam, most often achieve similar scores on other tests. The same can not be said of portfolios, peer assessment, and performance-based-assessment laboratory activities. Some students' marks vary widely among assignments. Teachers found this to be particularly true of performance-based-assessment activities. Doing well in one laboratory activity did not provide an indicator of future success. Teachers' concerns are supported by Hofstein (1988), who surveyed a number of different research projects and concluded that a strong correlation between practical laboratory exams and traditional tests does not exist. Performance-based assessment may not be supported by conventional testing.

The greater the teacher's acceptance of a traditional model of assessment, where reliability based on repeatability is used as a gauge for efficacy, the greater was the concern over adopting authentic assessment. Mary expressed the following concerns in one interview:

> We have tried a few performance-based-assessment activities. The one where we use the spec-20 to test the dilution and preparation of copper sulfate works well. But it raised many questions. Some of the better students have a lot of trouble and some of the weaker ones get good marks. It can inflate the marks of weaker students and drop the marks of better students. [pause] So we don't count them for very much.

When I asked Mary how she was classifying "weaker" and "stronger " students, she explained that her perceptions were based largely upon exam results. The quandary of not having other forms of assessment support testing was seen as a weakness. Mary used this same gauge to assign marks. A light meter on the spectronic-20 is used to measure concentrations of solution. Those with readings closest to the standard received the highest marks. Mary believed the assessment technique superior because it was objective. She expressed concerns over using more holistic marking schemes where multiple components were assessed simultaneously.

Marvin expressed a similar concern to Mary's with regard to portfolios. He was concerned that portfolio marks could be challenged because each category applied to so many different assignments and contained so many different components. For example, he stated concerns about assessing skill development as a single category because so many different skills were required in a single laboratory. Initially, he preferred to look at each skill individually. After attempting the approach, Marvin relented when he discovered the time required to accomplish the task was inordinately long.

Marvin also expressed the concern that portfolios might do a great deal to help weaker students but, for the most part, could only hurt those who did well on exams. Marvin explained that a 90% student (on the basis of testing) must get a 90% on the portfolio to maintain the mark, while a student who gets 40% on exams would most likely receive a passing grade because of portfolios. He felt that most students should get at least a 60% on portfolios. His greatest concern was that student work could be monitored during an exam but not in portfolios.

Marvin and Mary continue to see testing, based on a standardized assessment model, as objective, and therefore, superior. They believe that the superiority stems from the fact that test questions are constructed to match learner outcomes. During an action research meeting Mary explained that assessment can not be personalized to the manner in which individual learners have come to understand and construct knowledge. She explained that any assessment device has adverse effects on one type of learner.

To varying degrees, most teachers in the action research project took solace in knowing that knowledge, free from a subject, is superior because it is unbiased. All teachers in the project continue to use tests that have been constructed prior to meeting their students. Conversely, assessment that is individualized or linked with a subject was often deemed less valuable because it might contain bias. Decontextual knowledge was preferred over the idea that different students could be assessed in different ways to maximize their inherent strengths. Mary, Gary, Thad and Marvin indicated concerns about using performance-based assessment for group work. One of the major concerns was that the knowledge base for decision-making was elevated to the highest level within the group. Thad expressed the concern succinctly:

> One big drawback to using performance-based assessment for lab work is that students have to be placed in groups. Then you don't really know who understands. You know it like a lab group in my block four. [A student's name] knows what she is doing, but how about the rest of the group. I can tell you that [another student's name] has no idea what he is doing. Should he get a good mark because he can take orders?

Interviewer: But some students learn by talking. Why shouldn't they have some opportunity to be able to discuss things during an assessment situation. Scientists, for the most part work in groups. Cooperation and discussion are valued here.

Marvin: Yes, but we aren't training scientists. These kids have to be able to write an exam just to get into science.

Although teachers accepted the idea that students learn in different ways, there continued to be resistance toward allowing choices beyond those limited to portfolio assignments. The greatest portion of a student's report card mark continued to be assigned to formalized testing for all teachers interviewed at the end of the project.

Unanswered Questions

This action research study was not designed to establish cause-and-effect relationships between changes in assessment practices and teacher behavior. Rather it was designed to document changes within three schools as teachers moved toward changing their assessment procedures. Because so many factors were at play during the research project, schools were not compared to uncover reasons for success or failure. Like students, schools are highly idiosyncratic. The divergent ecologies of the differing schools help demonstrate that a movement toward the implementation of changing assessment practices can not be reduced to a repeatable methodology. A factor linked with success in one environment might have little consequence in another.

During the research study a number of teacher concerns or perceptions gave rise to questions that were not addressed. A list of potential questions is provided.

- Jane believes that females are more accepting of portfolio assignments than are males. She found that organizational skills, listening skills, and the ability to work in groups was generally less valued by males. The action research group did not collect data on the basis of gender, nor were questions of gender asked during interviews with either students or teachers. Might gender preferences in physics and other courses that have a lower number of females actually be a reflection of assessment strategies rather than orientation toward curriculum? A rationalist model is provided with greater support in an environment that uses mathematics to communicate, such as physics. Characteristically, English and Social studies rely less heavily on testing to determine student marks. These courses have characteristically had greater involvement by females (Alberta Education, 1992b). More authentic forms of assessing student work, such as writing assignments, group work, and special projects, have also experienced a longer history in courses identified with the fine arts or arts (Ashton-Warner, 1963; Collins, 1992; Eisner, 1985; Langer, 1957). Jane's hypothesis is strongly supported by the fact that more females currently take physics (1995) at Eastern High School than males. In 1991 only 31% of students taking physics at Eastern High School were female. Could the introduction of activities that are supported by more authentic forms of assessment, have influenced these changes?

- What are the long term effects of changing assessment strategies? At Eastern High School, English teachers and social studies teachers began using portfolios after discussions with science teachers. It is my perception that students, who now must complete portfolios in a number of courses, are less enthusiastic than they were in 1993/94. Another related question is whether or not the movement toward authentic assessment can act as a springboard to discussion between subject areas. Because a number of science teachers have assignments in two or more discipline areas, the opportunities for assessment discussions, along a broader base than science, have been initiated. Could the movement toward authentic assessment ever be instituted as a school-wide project?
- I worked as both researcher and facilitator within the action research groups. My dual role was not formally addressed during the studies. Was my role of researcher compromised by participation within the group. Yin (1988) cautions against active participation influencing interpretations. Did I see just what I wanted to see? Because the project took place over 18 months, more data was collected than used. Was the selection of interview quotes and the subsequent construction of survey questions influenced by my commitment to changing the assessment techniques used in my own classroom?
- Is it ever possible to escape an orientation of rationalism when the final message of assessment is conveyed to parents by way of a quantitative indicator on a report card? In the end, student records and descriptions of skill development are reduced to a number.

Given the delimitations of reporting student results, is it not better to develop assessment strategies that are totally supported by quantification?

Conclusion

The action research groups began to explore authentic assessment in February of 1993. By June of 1994 some substantial changes could be viewed. At all schools, teachers within the action research group indicated that they were more inclined to use the laboratory as part of their teaching. I personally viewed dramatic increases in laboratory usage at Eastern High School. I believe that the increased focus on students doing laboratory-related activities can partly be ascribed to the support derived from expanding assessment strategies; however, caution must be exercised in drawing simple cause-and-effect relationships. A number of other factors can be identified:

- Curriculum changes brought about a rethinking of teaching strategies to meet the goals of a curriculum based on a more inclusive notion of science that incorporated the Science-Technology-Society theme.
- The arrival of new textbooks to support changes in assessment.
- A new laboratory focus for junior high school a few years prior. These programs brought increased emphasis to doing science rather than reading about it. Because these students were now in high school, their expectations for a science course had changed.
- A provincial orientation toward accountability that was based on greater emphasis placed on assessment. Teachers knew that assessment would be given greater attention.

Specifically at Eastern High School, the increased laboratory emphasis could be linked with three very pragmatic changes in the school. These changes are:

- The influx of substantial cash that allowed the refurbishing of the science area.
- An addition of a full-time, experienced laboratory technician who not only encouraged laboratory activities, but also supported change by assisting teachers learn about new activities.
- The addition of three new members of the science department. These individuals brought new ideas and enthusiasm to the school.

The involvement of students in their own assessment brought about some positive changes in classroom climate. Survey results and interview discussions indicate that students want to have some choices. Students also expressed an understanding associated with the difficulties involved in arriving at fair assessment practices. Like teachers, students expressed frustration because they had been allowed few decisions in determining assessment practices. In addition, students identified many of the same problems with assessment practices that teachers did, when these practices were founded exclusively or dominated by testing. Students and teachers recognized that the precision of the report card mark was not confused with the meaning that mark represented.

It should not be surprising that a movement toward a more authentic assessment of learners is a slow process. The influences of rationalism extend far beyond a philosophical orientation that supports current structures. It influences what we interpret, and masks many taken-for-granted assumptions by providing a framework for interpreting reality. Rationalism employs logic to identify what is to be valued and what is unimportant. Questions of value and beliefs do not lend themselves well to discussions where rationalism is used as a way of understanding.

Rationalism is neither "good" nor "evil". It is a way of making sense of the world that we live in. Saul (1992) explains the allure of this methodology in that it permits prediction. Things that re-occur are considered facts, while events that defy prognostication are relegated a lower status. The appeal of structure can best be understood by the fact that it helps maintain power relationships. Within the field of education, this orientation helps define expertise and organizes a chain of authority from those who define the parameters of the test, to those who construct the test, and finally to those who must write the test.

The action research project was an attempt to begin asking questions about why current assessment practices exist. As a researcher I began by exposing the hidden and not so hidden assumptions. The movement toward portfolios allowed students to become involved in decision-making. Breaking down a structure that links the notion of "expert" to power and provides authority, requires that students and teachers begin viewing assessment in a different way. Assessment is not something that teachers do to students, but a way of coming to understand learning. Portfolios are an attempt to have students and teachers begin to participate in a process that gives value to the work of learners.

The use of performance-based-assessment arises from laboratory work, where process skills are acknowledged and group work is formally valued. Peer assessment and self-assessment grew from the laboratory component. The acceptance of performance-based-assessment strategies arose from teacher practical knowledge.

Teachers within the action research group drew from classroom experiences in stating that some students do not do well on exams. They explained that the very nature of exams masked knowledge that some students had gained, while inflating the apparent understanding of others. Poor reading ability, an inability to organize facts into cause and effect events, and low self-esteem were identified during action research meetings as factors that unduly influence exam writing abilities. Teachers also explained that some learners develop an understanding through discussions. The impetus for expanding assessment criteria, in the direction of performance-based-assessment, was an attempt to include a broader spectrum of tasks that might reflect abilities not previously recognized.

Although I stopped collecting any data after 18 months of research, the action research continues at Eastern and Northern High Schools in a less formal way. The dichotomy of curriculum intent and assessment strategies continues to exist. If anything the perceived gap is better understood and possibly wider. I believe that we have not resolved any issue with group consensus; however, that was never our goal. What I believe that we have accomplished, as action research groups, is a way of initiating discussions about teaching with colleagues and a way of including students in determining what they have learned.

Bibliography

Aikenhead, G. (1980). Science in social issues: Implications for teaching. Ottawa: Science Council of Canada.

Aikenhead, G. (1984). Teacher decision making: The case of Prairie High. Journal of research in science teaching, 21 (2), 167-186.

Alberta Education (1991). Teacher's resource manual for secondary science, Alberta Education.

Alberta, Education. (1992). A model for the assessment and evaluation of scientific problem solving (Draft No. 2, Alberta Education.

Alberta, Education. (1992b). Diploma Exam Review Report, Alberta Education.

Alberta, Education. (1993). Samples of students' responses: Grade 6. Alberta Education.

Alberta, Education. (1993). Samples of students' responses: Grade 9. Alberta Education.

Allan, K. A. & Miller, M. A. (1990). Teacher-researcher collaborative: Cooperative professional development. Theory into Practice, XXIX (Summer, Number 3), 196-200.

Anderson, J. (1989). Evaluating student achievement: Teacher practices and educational measurement. Alberta Journal of Educational Research, 35 (2), 123-134.

Aoki, T. (1978). Toward curriculum inquiry in a new key. In T. Aoki (Ed.), Phenomenological description: Potential for research in art education, . Concordia University: Concordia University.

Aoki, T. (1984). Interest, knowledge and evaluation: Alternative curriculum evaluation orientations. In T. Aoki (Eds.), Curriculum evaluation in a new key Edmonton: Secondary Education, University of Alberta.

Ashton-Warner, S. (1963). Teacher. New York: Simon and Schuster.

Baird, J. R. & Mitchell, I.J. (1986). Improving the quality of teaching and learning: An Australian case study – the PEEL project. Melborne: Monash University.

Berg, T. & Brouwer, W. (1991). Teacher awareness of student alternative conceptions about rotational motion and gravity. Journal of Research in Science Teaching, 28 (1), 3-18.

Berlak, H., Newman, F., Adams, E., Archibald, D., Burgess, T., Raven, J., and Romberg, T. (1992). Toward a new science of educational testing and assessment. Albany New York: State University of New York Press.

Bronowski, J. (1977). A sense of the future: Essays in natural philosophy. Cambridge Mass: MIT press.

Bronowski, J. (1981). Common sense of science. Cambridge Mass: MIT press.

Casonova, U. (1989). Research and practice: We can integrate them. Today, 7 (6), 44-49.

Chamber of Resources and Alberta Education. (1992). International comparison in education (Government of Alberta No. 1, Chamber of Resources & Alberta Education.

Cohen, M. & Harper, E. (1991). Student-as-scientist and scientist-as-student. Teaching education, 2 (3), 31-55.

Collins, A. (1992). Portfolio assessment for science education: Issues in purpose and structure. Science Education, 76 (4), 451-463.

Connelly, F., & Clandinin, D. (1985). Personal practical knowledge and modes of knowing: Relevance for teaching and learning. Chicago: University of Chicago Press.

Corey, S. (1990). Action research, fundamental research, and educational practice. In S. Kemmis and Mctaggart, R. (Eds.), The action research reader Victoria, Australia: Deacon Press.

Council of Ministers (1993) Information Bulletin for Schools: Science Assessment, School Achievement Indicators program, January 1996. Canada.

Council of Ministers (1994) Information Bulletin for Schools: Science Assessment, School Achievement Indicators program, January 1996. Canada.

Council of Ministers (1995) Information Bulletin for Schools: Science Assessment, School Achievement Indicators program, January 1996. Canada

Darling-Hammond, A. (1991). The Implications of Testing Policy for Quality and Equality. Phi Delta Kappan (November), 220-225.

Dennett, D. (1991). Consciousness explained. Boston: Little, Brown and Company.

Desmond, W. (1987). Desire, dialectic and otherness: An essay on origins. New Haven: Yale University Press.

Dimasio, A. (1994a). Descartes' Error and the Future of Human Life. Scientific American, 271 (October, Number 4), 144.

Dimasio, A. (1994b). Descartes' Error: Emotion, Reason, and the Human Brain (first ed.). New York: Grosset and Putnam.

Doran, L., Lawrencz, F., & Helgeson, S. (Ed.). (1994). Research on Assessment in Science. New York: MacMillan Publishing.

Dorr-Breme, D. (1983). Assessing students: Teacher routine practices and reasoning. Washington: National Institute of Education, ERIC document reproduction No Ed 242-742.

Douglas, C. & Kahle, J. (1983). A profile of NABT: Results of 1982 national survey. American Biology Teacher, 45 (8), 410-414.

Duffee, L., & Aikenhead, G. (1992). Curriculum change, student evaluation, and teacher practical knowledge. Science Education, 76 (5), 493-506.

Eisner, E. (1985). The art of educational evaluation. Philadelphia: Falmer Press.

Eisner, E. (1993). Reshaping assessment in education: some criteria in search of practice. Journal of Curriculum Studies, 25 (No. 3), 219-233.

Elliot, J. (1991). Action research for educational change. Milton Keynes, England: Open University press.

Erickson, G. et. al (1993). British Columbia: Assessment of Science 1991 (Technical report II: Student performance component No. FCG 187). Ministry of Education Province of British Columbia, Victoria.

Feimer-Nesmer (1986). The culture of teaching. In M. C. Wittracks (Eds.), Handbook on research on teaching New York: MacMillan.

Feyerbend, P. K. (1975). Against method: Outline of anarchistic theory of knowledge. London: New Left Books.

Gadamer, G. (1993). Truth and method (Mohr, P., Trans.). (second ed.). New York: Continuum.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Geddis, A. (1991). Improving the quality of science classroom discourse on controversial issues. *Science Education*, *75* (2), 169-183.

Goodlad, J. (1983). A study of schooling: Some Findings and hypotheses. *Phi Delta Kappan*, *64* (7), 486-501.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Groome, T. H. (1980). *Christian religious philosophy*. NewYork: Harper Collins.

Gunstone, R. (1988). Learning in science education. In P. Fensham (Eds.), *Developments and dilemmas in science education* London: Falmer Press.

Herman, J. (1992a). What research tells us about good assessment. *Educational Leadership* (May), 74-78.

Herman, J. (1992b). A practical guide to alternative assessment. *ASCD*.

Herman, J., & Golan, S. (1992). *Effects of standardized testing on teachers and learners* No. CSE Technical Report 334). National Center for Research on Evaluation, Standards, and Student Testing, University fo Calfornia.

Hofstein, A. (1988). Practical work and science education II. In P. Fnesham (Eds.), *Development and dilemmas in science education* (pp. 169-189). London: Falmer Press.

Jackson, P. (1968). *Life in classrooms*. New York: Holt Rhinehart Winston.

Kemmis, S. & McTaggart, R. (1990). *The action research planner* (3 ed.). Victoria Australia: Deakin University press.

Klainin, S. (1988). Practical work and science education I. In P. Fensham (Eds.), *Development and dilemmas in science education* (pp. 73). London: Falmer Press.

Kuhn, T. (1970). *The structure of the scientific revolution*. Chicago: University of Chicago Press.

Lakatos, I. (1963). Proofs and refutations. *Britsih Journal for the Philosophy of Science*, *14*, 1-25.

Lakatos, I., & Musgrove A. (1974). *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Langer, S. (1957). *Problems of Art*. New York: Charles Scribner and Sons.

Lantz, O., & Kass, H. (1987). Chemistry teacher's functional paradigms. *Science Education*, *71* (1), 117-134.

Levine, D. (1985). *The Flight from Ambiguity*. Chicago, Illinois: University of Chicago Press.

Lewis, A. (1992). National Standards. *Education Digest*, Jan 1992, Vol 57, Issue 5, p. 53-55

Lindberg, D. (1992). *The beginnings of Western Science*. Chicago: University of Chicago Press.

Madaus, G., & Kellaghan, A. (1989). The British Experience with "Authentic" Testing. *Phi Delta Kappan* (February), 459-469.

Maeroff, G. (1991). Assessing alternative assessment. *Phi Delta Kappan* (12), 272-281.

Martin, B. & Brouwer, W.. (1991). The sharing of personal science and the narrative elements in science education. *Science Education*, *75* (6), 707 - 722.

Maturana, H., & Varela, F. (1987). *The tree of knowledge*. Boston: Shambhala.

McLean, L. D. (1985). The craft of student evaluation in Canada: A report to the C.E.A. on policies, practices, and the uses of assessment on achievement No. ED 251 159). Canadian Education Association.

Medawar, P. (1983). Limits of Science . London: Metheun.

Mehrens, W. (1992). Using performance assessment of accountability purposes. Education Measurement: Issues and Practices (Spring), 3-10.

Midwood, D., O'Connor, K., & Simpson, M. (1993). Assess for success: Assessment, evaluation, and reporting for successful learning , Ontario Secondary Teachers' Federation, Toronto Ontario.

Mielnichuk, J., & Peat, D. (1990). Teaching thinking: Enhancing learning (Government No. Alberta Education, Curriculum Branch.

Miles, M. B. Huberman, A.M. (1984). Drawing valid meaning from qualitative data: Toward shared craft. Educational Researcher , 13 (5), 20-30.

New Brunswick. (1992). Biology Curriculum Guide 2201, 2202 (Curiculum Guide No. New Brunswick Department of Education.

Newfoundland (1990). Biology curriculum guide (Curriculum guide No. Newfoundland Department of Education.

Nova, S. (1993). Biology curriculum guide (Draft No. Ministry of Education, Province of Nova Scotia.

Nuttal, D. (1992). Performance Assessment: The Message from England. Educational Leadership (May), 54 - 59.

O'Neil, J. (1992). Putting assessment to the test. Educational Leadership (May), 14-19.

Ontario Minstry of Education. (1981). Onatrio assessment instrument pool - Physics , Ontario Ministry of Education, Toronto.

Ontario Minstry of Education. (1983). Onatrio assessment instrument pool - Chemistry , Ontario Ministry of Education, Toronto.

Ontario Minstry of Education. (1987). Curriculum guidelines: Science, intermediate and senior divisions (Curriculum Guide No. Ontario Ministry of Education.

Ontario Minstry of Education. (1989). Onatrio assessment instrument pool - Biology , Ontario Ministry of Education, Toronto.

Orlich, D., et. al. (1980). Teaching strategies: A guide to better instruction . Lexington, Mass.: D.C. Heath.

Perrone, V. (1978). Supporting teacher growth. Childhood Education , 54 (6), 298-302.

Pipho, C. (1989). Outcomes or 'Edubabble'. Phi Delta Kappan (May), 662-663.

Polanyi m. (1962). Personal Knowledge (second ed.). Chicago: University of Chicago Press.

Popham, W. J. (1993). Circumventing the high costs of authentic assessment. Phi Delta Kappan (February), 470.

Province of British Columbia (1986). British Columbia science assessment 1986 - Summary report, Victoria.

Quebec Education. (1987). Curriculum guide: Secondary schools human biology (Curriculum Guide No. Province of Quebec.

Risi M. (1982). Macrosole: A holistic approach to science teaching . Ottawa: Science Council of Canada.

Ritter, R. (1991). Science technology and society. In A. Education (Eds.), *Teacher resource manual for high school science* (pp. 75-82). Edmonton Alberta: Alberta Education.

Ritter, R., & Samiroden, W. (1993). *Nelson Biology: Teacher's Edition*. Scarborough, Ontario: Nelson Canada.

Rodgers, T. (1992). *Principles for fair assessment practices for education in Canada* (Centre for Research and Applied Measurement in Education), Joint advisory Committee, University of Alberta.

Ryan, C. D. (1994). *Authentic assessment*. Westminster, California: Teacher Created Materials, Inc.

Sanger, J. (1990). Awakening a scream of consciouness: The critical group in action research. *Theory into Practice*, *XXIX* (number 3, summer), 174-178.

Saul, J. R. (1992). *Voltaire's Bastards*. New York: Penguin Books.

Schon, D. (1983). *The reflective practitioner*. London: Jossey-Bass.

Schwab, J. J. (1969). *College curricula and student protest*. Chicago: University of Chicago press.

Science C.ouncil of Canada. (1984). *Science for every student: Educating Canadians for tomorrow's world* No. 36). Science Council of Canada, Government of Canada.

Shavelson, R., & Baxter, G. (1992). What we've learned about assessing hands-on science. *Educational Leadership*, *75* (May), 20-24.

Shumsky, A. (1990). Cooperation in action research: A rationale. In S. a. M. Kemmis R. (Eds.), *The Action Research Reader* (pp. 81-85). Victoria, Australia: Deakin University.

Simmons, K., Kobett, B., Hess, J., and Fennell, F. (1992). Linking instruction and assessment in the mathematics classroom. *Mathematics Teacher* (Feb), 11-17.

Smith, D. (1988). Brighter than a thousand suns: Facing pedagogy in the nuclear shadow. In T. Carson (Eds.), *Rethinking and Reorienting Curriculum* Edmonton, Alberta: World Council for Curriculum and Instruction.

Smith, M. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, *20* (5), 20-24.

Stake, R. (1972). The countenance of educational evaluation. In T. a. Cowley (Eds.), *Knowing in curriculum evaluation* Dubuque, Iowa: W.C. Brown.

Tierney, R., Carter, R., and Desar,L. (1991). *Portfolio assessment in reading classrooms*. Norwood MA: Christopher Gordon.

Timpane, J. (1995). How to convince a reluctant scientist. *Scientific American*, *272* (January, #1), 104.

Tobin, K., Espinet, M., Byrd, S, & Adams, D. (1988). Alternative perspectives of effective science teaching. *Science Education*, *72* (4), 433-451.

Vavrus, L. (1990). Put portfolios to the test. *Instructor*, *100* (1), 48-51.

Villani, A. (1992). Conceptual changes in science and science education. *Science education*, *76* (6), 223-237.

White, R. (1988). Theory into practice. In P. Fensham (Eds.), *Developments and dilemmas in science education* New York: Falmer Press.

Wiggins, G. (1989). A true test: Toward more authentic assessment. *Phi Delta Kappan*, *70* (9), 703-713.

Wiggins, G. (1992a). Creating test worth taking. Educational Leadership (May), 26-33.

Wiggins, G. (1992b). Making Sense of Authentic Assessment. Instructor, 102 (Nov 92), 40-42.

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. Phi Delta Kappan, 75 (3), 200-215.

Wilson, R. (1989). Evaluating student achievement in Ontario high schools. The Alberta Journal of Educational Research, 35 (2), 134-144.

Wolf, D., Bixby, J., Glenn, J. and Gardner, H. (Ed.). (1991). To use their minds well: Investigating new forms of student assessment. Washington D.C.: American Educational Research Association.

Yin, R. (1988). Case study research: Design and Methods. Newbury park, California: Sage Publication.

Zielder, D., Lederman, N., Taylor S. (1991). Fallacies and student discourse: Conceptualizing the role of critical thinking in science education. Science Education, 76 (4), 437-450.

Thesis Appendices

Appendix #1: Faculty of Education Ethics Review

Project Title: Exploring Alternate Assessment Strategies for High School Science.
Applicant(s): Bob Ritter
Department(s): Secondary Education. Date: Jan. 8 /.93

Statement of Problem and Methods:
The 1992 implementation of a new science curriculum calls for an STS (science, technology, and society) approach to teaching science. As teachers struggle to come up with new teaching strategies to meet these goals, a perceived division exists between the manner in which students are evaluated and the manner in which curriculum is intended to be presented. The following questions have surfaced during meetings with teachers:

1. Are evaluation strategies being used by high school science teachers to assess students' knowledge of science congruent with the intents of the new curriculum? Have assessment approaches changed?
2. Can we develop assessment strategies that complement STS teaching strategies? What components can be evaluated? What message will we send to students if we don't evaluate a certain component of STS ? Will the students value components of the course not evaluated ?
3. What problems will be created as we attempt to move toward assessing technological thinking and social issue components of a science program?

The study will be conducted as an action research project in which teachers from five high schools work collaboratively to develop and implement alternate assessment strategies. Data will be collected from selected teachers within the action research group by way of interview. Transcripts will be developed from the interview and subject to teacher approval. Interpretations of interview data will be developed with interviewee and subject to approval by the interviewee. I also intend to keep a researcher journal outlining my impressions of the progress or difficulties experienced by the action research group.

Who are the participants and how will they be involved in your research?
Teachers from XXX requested that we explore the congruency of curriculum intent and student assessment. Selected Teachers from Southern and Eastern also expressed an interest in the action research. Principals at Southern and Eastern, as well as Muriel Dunnigan, Deputy Superintendent for Edmonton Catholic Schools, expressed support. Eight students from Eastern will be asked to provide their impressions by way of interview. In all situations volunteers will be requested.

How will the nature and purpose of the research be explained to the participants.
Teachers from the Northern High School identified a need for the study. Some teachers from Southern and Eastern High Schools indicated that they were interested in becoming involved in the action research group. As an action research group, the nature of the problems investigated and the methodology used must be mutually agreed upon by the research group. An outline of research methodology has also been prepared for teachers who will be involved in the interview process. The research proposal will also be supplied to senior administration for school boards concerned. Selected participants will be asked to provide two 20 minute interview about their perceptions of alternate assessment strategies developed by the action research group. Four students at Eastern will be asked to volunteer for a group interview to provide their impressions of alternate assessment strategies.

How will informed consent of participants be obtained?
Teacher interviewees will be provided with informed consent forms. Transcripts will maintain the anonymity of the teachers involved and teachers will be asked to check validity of interpretations. Pseudonyms will appear in the transcripts. Student interviews will be restricted to students at Eastern High School and parental permission will be required. Students rights will parallel those of the teachers involved in the study. The research proposal will be made available to parents and students. Students and teachers will have the right to withdraw at any time from the study.

Are captive or dependent populations used?
No, only volunteers will be used for interviewing. Teachers will decide which alternate assessment strategies are appropriate for their classroom.

How will provision be made for exercising the right to opt out at any time?
Students are only requested to provide a single interview. Volunteers are sought and parental permission must be given. Teachers can also opt out of the study at any time. Teachers are not required to implement any set of strategies but work within an action research group to develop STS assessment strategies.

How will confidentiality and anonymity be maintained?
Teacher and student pseudonyms will be used for the transcripts. Teachers and students will be provided with interpretations for validation.

Is deception and/or risk involved in the project? Is so, how will the interests of the subjects be protected?
Deception is not employed.

Appendix #2: Cooperative Activities Program

Field Experiences	
Faculty of Education	392-3661
University of Alberta	T6G 2G5

---

1. **Nature of the Activity**
   Course related
2. **Organization to be involved**
   St. Albert Protestant/ Separate Edmonton Catholic School Board
   upon request.
3. **Requester:**
   Dr. Wallie Samiroden Faculty of Education
   Professor Secondary Education
   Room 348 492-3676
   On behalf of: Robert Ritter
   Room 333 Ed. South 492-1731
   430-6794
4. **Description of Activity:**

Title: Exploring alternate assessment strategies for high school science.

Objectives:

- To work with teachers in an action research group to develop assessment strategies that support teaching science within an STS (science, technology, society) context.
- To determine what problems will be created as we attempt to move toward assessing technological thinking and social issue components of a science program?
- To explore teachers' reasons, within the framework of teacher practical knowledge, for selecting the evaluation practices that they use to evaluate student knowledge?
- To explore teachers' reasons, within the framework of teacher practical knowledge, for including or excluding assessment practices that could evaluate student knowledge of the three STS components: nature-of-science, science-and-technology, and science-and-society?
- To bridge the gap between curriculum intent and evaluation practices (or should we attempt to bridge the gap)? and to determine if there is a shared meaning and language for STS?
- To identify perceived constraints associated with the institutional setting (because of pressures from the School District, Alberta Education, or the community) inconsistent with developing aspects of STS curriculum and evaluation?

Procedures:

- A subgroup of teachers will be selected for the research. The study will be conducted at two Edmonton Catholic schools, Southern and Eastern and Northern High School. At least one teacher from each school will be selected from the volunteers for in-depth interviews. Teachers involved as interviewees will be asked to keep a log book of their impressions about how new assessment strategies are perceived by students, their colleagues, and themselves. Their impressions, recorded in point form, will act as an outline to construct interview questions. Interviews will be transcribed, coded, and analyzed. Follow-up interviews and validation of interpretations will be conducted with each interviewee. (Teacher permission will be sought for all interpretations). Teacher identity will be protected, if requested by the teacher or school board.

- Selected student interviews will be conducted as new assessment strategies are implemented. (The possibility of involving student teachers also exists). Student identity will be protected and parental permission for each interview will be sought.
- A series of organizational meetings have been set up with all interested teachers at the five high schools involved. A common focus has been established: "developing assessment strategies that support the teaching of science within an STS context". Working groups, at each of the high schools, will begin developing alternate assessment strategies. A list of possibilities has been provided at exploration meetings. Larger group sharing among schools has also been requested, but no teacher is asked to follow a master plan for assessment. The alternate approaches to assessment will be developed independently and ideas for assessment will be shared among schools involved.
- Researcher journals will also be used. The journal will record my impressions of the changes seen in my own classroom setting as alternate assessment strategies are implemented. The journal will also serve as a record of my impressions from organizational meetings.

Evaluation techniques:

- Analysis and coding of interview transcripts. Teachers and student interviewees will be asked to validate interpretations.
- Researcher journal examined and interpretations scrutinized by colleague.

5. Anticipated value to school(s) or school district(s) involved.

- Teachers have the opportunity to work collaboratively to develop alternate assessment strategies that support new directions in curriculum development.
- Students will have an opportunity to participate in formulating an assessment strategy that more accurately reflects curriculum intent and student needs.

6. Suggested personnel, schools and times.

Schools: Northern (request from Robert Hogg); Southern and Eastern (contact and support from Muriel Dunnigan, Gerry Wowk, Merv Lynch and the school principals).
Times: February 1 93 to April 30 /95.

Appendix #3 Parent approval letter

March 18/93

Dear Parent:

The purpose of this letter is to seek your permission to discuss alternative assessment strategies with your daughter/son. The discussion will serve as the foundation for information gathering toward new directions in assessment practices for science classrooms, a topic I am pursuing for my Ph.D. degree. Although laboratory skills, scientific discourse, and the development of problem-solving strategies are emphasized during classroom activities, what continues to be most valued is factual knowledge, as determined by traditional exams. A movement toward alternative assessment strategies is not an attempt to move away from exams, but rather recognize other components of classroom activities that provide a more complete picture of student learning.

The commitment on the part of your daughter/son would consist of two 20 minute interviews providing feedback on alternative assessment strategies being employed in their science course or courses. The interview would be audio-taped and transcribed for data toward my thesis. The student's name would be changed to maintain anonymity. Selected interviews would only be done on a volunteer basis, and at a time convenient for your daughter/son.

The seeking of parental permission is but one of many steps that I have taken to ensure the learning experiences of students are not compromised. My proposal has been adjudicated by and received approval from an ethics review committee, in the Department of Secondary Education, at the University of Alberta, and another committee from Edmonton Catholic Schools. I believe that student input into assessment is critical to changing the way in which things are valued in science classrooms.

Thank you for your cooperation.

Parent signature of approval

Appendix #4: Sample student interview questions

Interviews carried out in groups of two or three. Interview questions may be phrased differently and presented in a different order depending upon responses. Prior to the interview students will be asked to reflect upon their experiences with alternative assessment. I will explain that any rejection of different ways of assessment does not indicate any disagreement with either my approach or that of their student teacher, the person involved in marking most of the grade 11 material work from February 1993 to April 15, 1993.

1. Do you believe that the portfolios were worthwhile?
    - Please indicate some positive and negative aspects about doing portfolios.
    - Are exams a better way to assess student learning? Explain.
    - Should students be given some choices in what is to be marked, as per their portfolios?
    - What do you believe that your teacher believes is most important (exams or portfolios).
    - In your opinion what is the reason for using portfolios?
    - Should teachers mark every assignment and every question that students submit?
    - Should all of the work that students do be evaluated?

2. Are the metacognitive journals worthwhile?

- Please indicate some positive and negative aspects about doing the journals.
- Have you formulated any questions on your own from doing journals.
- Have you done any concept maps on your own?
- Have you discussed journal entries with anyone else in the class? Have you read someone else's journal entry? Has someone else read one of yours?

3. Can you learn chemistry from writing a murder mystery?
    - Please indicate some positive and negative aspects about doing the murder mystery.
    - How do you feel about peer assessment for the murder mystery?

Appendix #5: Students' Questionnaire Portfolio Assessment

Use the computer scoring cards to provide your opinions about the following statements.
A = agree, B = agree most of the time, C = neutral response (neither strongly agree or disagree), D = disagree most of the time, E = disagree.

Descriptor	%A	%B	%C	%D	%E
1. Portfolio assessment is worthwhile.	52.1	22.5	16.2	1.4	7.1
2. I believe that students should have some choices in determining which assignments should be marked.	40.9	23.9	23.9	5.6	4.2
3. I prefer individual lab write ups that provide immediate feedback.	16.2	24.7	34.5	15.5	8.5
4. I prefer a seminar after the laboratory activity or case study where we have an opportunity to listen to the answers given by other students.	35.9	27.4	21.8	5.6	9.2
5. Portfolios require less work than individual laboratory write ups and hand in questions.	4.9	8.5	19.0	28.9	37.2
6. Everyone should hand in the same material to be marked in order to get a fair assessment.	29.6	22.5	18.3	13.4	15.5
7. Portfolios serve as a good source of review.	43.7	28.9	13.4	7.0	5.6
8. We shouldn't use portfolios because it only provides a sample of our work. All of our work should be marked.	3.5	6.3	19.0	35.9	33.1
9. I find it time consuming to select work for the portfolio.	14.1	25.4	38.7	9.9	10.6
10. I tend to read over labs, case studies, questions, and journal entries prior to handing in the portfolio.	36.6	33.1	14.1	6.3	8.5
11. I believe the marking scheme used for the portfolio is fair.	37.3	36.6	14.1	7.0	2.1
12. I believe that the portfolio mark will lower my course mark.	14.1	10.6	26.1	19.1	30.0
13. I believe that the marks for the portfolio reflect those of the exams. Students who get high exam marks do well on portfolios while those who do poorly on exams always do poorly on portfolios.	9.2	15.5	27.5	23.9	22.5
14. Individuals who always do their work, but have trouble writing exams benefit by portfolios.	53.5	29.6	12.7	2.1	1.4
15. Portfolios create greater opportunities for students to copy the work of other students than assignments which are handed in shortly after they are completed (the next day or two).	15.5	23.9	31.0	16.9	12.0
16. I believe that journal entries that include concept maps, challenge questions, and my personal reactions to what I have learned have been useful.	26.8	31.7	26.7	6.3	7.0
17. Journal entries have no place in the portfolios because they have no single right or wrong answer.	7.0	13.4	21.1	29.6	26.8
18. I believe that exams provide a better indication of what students have learned in the course than do portfolios.	14.1	18.3	31.7	16.2	19.0
19. The teacher should choose all elements that go into the portfolio.	9.9	11.3	29.6	20.4	27.5
20. The teacher should mark everything in the portfolio. Sample marking will not provide a clear picture.	20.4	19.7	38.7	10.6	9.2

Appendix #6: Assessment Templates

6A: Assessment template for Peer Assessment (group)

Peer assessment of another group

Evaluator's Name:______________________________

Group members assessed____

Use the following rating scale: #4 is excellent, #3 is good, #2 is average, and #1 is poor. Circle the appropriate number

1. Laboratory design was thoughtful	1	2	3	4

Comment:

2. Data was organized and communicated effectively.	1	2	3	4

Comment:

3. Conclusions were supported by the data collected.	1	2	3	4

Comment:

6B: Assessment Template for Group Work

Individual assessment of the group

Evaluator's Name:________________________________

Group members ____

Activity name :________________________________

Use the following rating scale: #4 is excellent, #3 is good, #2 is average, and #1 is poor. Circle the appropriate number

1. The group worked cooperatively. Everyone assumed a task and carried it out. (You are not asked to identify people who did not contribute). Everyone's opinion was valued.	1	2	3	4

Comment:

2. The group was organized. Materials were collected for samples, and problems were addressed as a group.	1	2	3	4

Comment:

If you were to repeat the experiment what things would you change? (How would you improve your design.

6C: Assessment Template for Debates

Category	Value		
• Evidence of research in preparation for the debate	1	2	3
comments:			
• Points expressed clearly and concisely.	1	2	3
comments:			
• Scientific and technological accuracy are demonstrated	1	2	3
comments:			
• Listens to others and responds to their arguments in a positive manner.	1	2	3
comments:			
• Demonstrates tolerance of alternate viewpoints	1	2	3
comments:			
• Supports his or her point of view when questioned with well-thought out responses.			
comments:			

6D: Assessment Template for Presentation of Research Topics

Note: Depending upon the nature of the research question, some of the categories may not be appropriate.

Category	Value		
• Sources used during research are documented.	1	2	3
comments:			
• Information is expressed clearly and concisely.	1	2	3
comments:			
• Scientific and technological accuracy are demonstrated	1	2	3
comments:			
• Social implications / environmental implications are considered in presentation.	1	2	3
comments:			
• Risk / benefit analysis is presented.	1	2	3
comments:			
• Limitations of scientific and technological approaches are acknowledged.			
comments:			

6E: Assessment Template for a Teacher-Directed Laboratory Investigation

Value	Laboratory Skills	Participation	Write-up (product)
5	• Understands the reason for the investigation • Can identify variables • Demonstrates safety in carrying out procedures. • Evaluates appropriateness of methods used. • Lab area is left clean.	• Gets involved quickly • Stays involved throughout the investigation. • Provides leadership for the lab group	• Write-up is organized and easy to follow. • Observations are insightful, and accurate. • All questions are completed. • Evidence of further research, lateral thinking, or critical thinking.
4	One of the components is missing or poorly developed.	• Gets involved quickly • Stays involved throughout the investigation. • Leadership is not demonstrated.	• Work is complete and accurate but no evidence of lateral thinking, critical thinking or further research.
3	Two components are missing	Needs periodic reminding to stay on task	• Components of the write-up are missing. or • Some inaccuracies appear in write-up. • Acceptable presentation but organization needs work.
2	Multiple errors prevented the individual / group from obtaining any reliable data	Needs constant reminder to remain on task.	Incomplete or Multiple error
1	Investigation is completed but write-up is not done		
0	Investigation is not done		

Comments:

6F: Assessment Template for a Student-Designed Laboratory Investigation

Value	Laboratory Skills	Participation	Write-up (product)
5	• Is able to formulate a hypothesis or state the reason for the investigation. • Can identify variables • Design is insightful • Procedure is safe. • Evaluates appropriateness of methods used. • Lab area is left clean.	• Gets involved quickly • Stays involved throughout the investigation. • Provides leadership for the lab group	• Data presentation by tables and graphs are organized and easy to follow. • Observations are insightful and accurate. • Conclusions are supported by data. • Evidence of further research, lateral thinking, or critical thinking.
4	One of the components is missing or poorly developed.	• Gets involved quickly • Stays involved throughout the investigation. • Leadership is not demonstrated.	• One of the components described above is missing.
3	Two or more components are missing.	Needs periodic reminding to stay on task	• Components of the write-up are missing. or • Some inaccuracies appear in write-up. • Acceptable presentation but organization needs work.
2	Multiple errors prevented the individual / group from obtaining any reliable data	Needs constant reminder to remain on task.	Incomplete or Multiple error
1	Investigation is completed but write-up is not done		
0	Investigation is not done		

Comments:

6G: Assessment Template for Formative Assessment of Student Problem-Solving

Function: formative assessment

Date________________________

Students name ______________

Laboratory Activity________________________________

Check list (√)	Attitude or Skill Demonstrated	Comment
1.	Likes to solve problems	
2.	Preservers -sticks with problem-solving	
3.	Willingness to share ideas with others	
4.	Understands the question being asked	
5.	Is able to formulate a plan	
6.	Considers alternatives and demonstrates flexibility in attempting other solutions	
7.	Can synthesize data or eliminate data during problem-solving	
8.	Checks data and / or approach for accuracy	
9.	Can develop a conclusion	
10.	Effective communication of problem-solving strategy and results.	

6H: Assessment Template for Scientific Inquiry

Function: formative assessment

Date________________________

Student's name ______________________

Laboratory Activity__________________________________

Place a check (√) in the appropriate box.

Quality	frequently	sometimes	never
Selects appropriate strategies			
Accurately implements strategies used for solutions			
Self-initiates reflection of strategy and considers alternate strategies			
Approaches scientific inquiry in an organized systematic manner.			
Integrates library research skills with laboratory-based inquiry.			
Listens to alternate ways of solving problems even after a solution is found.			

6I: Assessment Template for Biological Drawings

Use the following check list:

Descriptor	Yes	No
1, Use of blank paper and a pencil for diagrams		
2. Drawing at least 0.25 of page.		
3. Title provided for diagram.		
4. Magnification indicated and scale of the diagram is present.		
5. No shading or coloring.		
6. Proportions resemble that of cell tissue or organ.		
7. Detail indicates that only visible structures have been presented.		
8. Labeling is correct.		
9. Labels are parallel and to the right of the diagram.		
10. Diagram is neat.		

6J: Portfolio Assessment: Science 10 Unit 2, 1992

Your portfolio is worth 30% of your term mark, make sure to have all components of the portfolio well organized. The portfolio provides a description of your work as a student of biology. It is a combination of work that you have chosen and your teacher has set as a requirement.

- It is recommended that portfolios be organized in Duo Tangs or three ring binders.
- Make sure to include a table of contents at the beginning of your portfolio. You may decide what goes where. There is no one correct way!
- Number all pages in your portfolio.
- On pages where you have answered multiple questions, but have selected specific questions to demonstrate your learning make sure that you identify the questions to be used for assessment.

1. Laboratory work: 7 activities required.
Three activities have been identified as compulsory. Choose four other activities must be selected from the optional list below.

Title	Compulsory	Optional
Investigation 10.3: Working with the Microscope		√
Investigation 10.4: Using the Microscope		√
Case Study 10.6 Life from Non Living Things	√	
Investigation 10.5: Using the microscope for forensic investigations		√
Investigation 10.8: Comparing Plant and Animal Cells	√	
Investigation 10.11: Structure and Function of Cells		√
Investigation 11.3: Observing Diffusion and Osmosis	√	
Investigation 11.4: Discovering the Effects of Osmosis on Cells		√
Investigation 11.6: Observing Living Organisms in Pond Water		√
Investigation 11.7: Responses of Paramecium		√
Investigation 11.9: Are Bacteria Found in Milk?		√
Case study 11.10: Monitoring bacteria Levels in Mascara		√
Investigation 12.3: Observing Cell Division	√	
Investigation 12.4: Determining the Rate of Cell Division		√

2. Select 20 questions from the following work sheets, or written assignments. The questions should indicate originality, creativity, and commitment to learning biology. One question need not be chosen from every assignment. Some students may choose a number of questions from more challenging assignments. If you did any extra reading or spoke with anyone to help you provide a complete explanation, please footnote in your portfolio. Discussions or readings beyond the confines are good things to do. Make sure that you provide an indication of which questions you have chosen. The symbol (√) is used to designate compulsory questions. The symbol (◊) indicates optional questions.

√ Problem solving question # 14 from chapter reviews of chapter 10
√ Problem solving question #12 from chapter reviews of chapter 11
◊ Any of the Self Check Questions from chapters 10, 11, and 12
◊ Any of the Review Questions from chapter reviews of chapter 10, 11, and 12

- ◊ Any of the problem solving questions from chapter reviews of chapters 10, 11, and 12
- ◊ Any of the critical thinking questions from chapter reviews of chapters 10, 11, and 12
- ◊ A single completion question from unit review
- ◊ A single true / false question from unit review
- ◊ A single multiple choice question from unit review
- ◊ A single short answer question from unit review
- ◊ The matching activity from the unit review
- ◊ The crossword puzzle from the unit review.
- ◊ Any of the Challenge questions from the unit review
- ◊ Any of the problem-Solve questions from the unit review

3. Getting Started Entry and Ask Yourself entries are all compulsory.

4. Select two "Challenge" or "Try This" activities from any of chapters 10, 11, or 12

6K: Scoring Rubric for Portfolios: Science 10 (unit 2), 1992

UNIT 2 CELLS

• Development of scientific skills	5 marks
• Identifies problem and is capable of framing question. • Rules for safety are observed when the student devises a procedure. • Student identifies discrepancies in data during discussions with laboratory group and some consultation with teacher. Student assumes responsibilities for initiating discussions about scientific inaccuracies. • Identifies cause and effect relationships, resulting from data collection or quantitative analysis. • Identifies and uses independent and dependent variables for investigations. • Proper use of microscope, wet mounts are prepared, and magnification of the image is calculated. • Rates of cell division and growth can be calculated.	
• Scientific accuracy	5 marks
• Scientific terminology is used appropriately. • Student can calculate the magnification of an image viewed under the microscope. • Student looks for trend or relationships in developing scientific theories and/or laws. Can differentiate between living and nonliving things, plant and animal cells, and normal and cancerous cells. • Student can identify parts of a cell and state the function of each. • Student can outline the levels of cellular organization in the human body. • Explains the problem investigated. Provides explanations and draws conclusions related to hypothesis. Modifies theory and/or hypothesis on the basis of results.	
• Communication	5 marks
• Portfolio is well organized with a table of contents and identified questions and answers. • Organizes and presents data in tables and graphs. • Scientific drawings are presented in an acceptable manner. • The student's ideas are expressed in a purposeful and coherent manner. The focus is maintained throughout the presentation and transitions connect ideas appropriately. • Student initiated reconnaissance of experimental design and/or data collection. Student discusses the limitations of the experimental design or data collected. Evaluates assumptions and the effects of bias. Evaluates experimental design and re-structures experiment.	
• Conventions	5 marks
• Writing is essentially free from error in spelling, punctuation, and grammar. Errors that are present do not affect the clarity of presentation. • Rules for graphing are followed. • Rules for scientific drawings are followed. • Proper units are chosen for measurement. SI metric rules are followed. • Bibliography is included for scientific reports or research assignments.	

6L: Sample Portfolio Assignment: Biology 30, Term #1, 1992

Your portfolio is worth 30% of your term mark, make sure to have all components of the portfolio in place and well organized. The portfolio provides a description of your work as a student of biology. It is a combination of work that you have chosen and your teacher has set as a requirement.

- It is recommended that portfolios be organized in Duo Tangs or three ring binders.
- Make sure to include a table of contents at the beginning of your portfolio. You may decide what goes where. There is no one correct way!
- Number all pages in your portfolio.
- On pages where you have answered multiple questions, but have selected specific questions to demonstrate your learning make sure that you identify the questions to be used for assessment.

What is submitted:

The symbol(√) indicates compulsory, while the symbol (◊) indicates that an optional selections may come from this area.

1. Laboratory work: One laboratory has been identified as compulsory (The effects of pH on protein digestion). Three other labs must be selected from the list below.
    √ Effects of pH on protein digestion (compulsory)
    Choose three from the following list:
    ◊ Passive transport
    ◊ Enzymes and H_2O_2
    ◊ Identification of carbohydrates
    ◊ Identification of proteins and lipids
    ◊ Case study: Control of digestion
2. Select 10 questions from the following work sheets, or written assignments. The questions should indicate originality, creativity, and commitment to learning biology. One question need not be chosen from every assignment. Some students may choose a number of questions from more challenging assignments. If you did any extra reading or spoke with anyone to help you provide a complete explanation, please footnote this in your portfolio. Discussions or readings beyond the confines are good things to do. Make sure that you provide an indication of which questions you have chosen.
    ◊ Question sheet on passive transport.
    ◊ Enzymes questions sheet.
    ◊ Biochemistry review questions (page 172).
    ◊ Biochemistry critical thinking questions (page 175).
    ◊ Digestion diagram
    ◊ Digestion handout sheet.
    ◊ Digestion application questions
    ◊ Digestion critical review questions.
3. Journal entries: all are compulsory.
    √ Description of a question that may have given you difficulty from the exam on cell processes.
    √ Concept map or maps from biochemistry.

6M: Scoring Rubric for Portfolios: Biology 30, 1992

• Development of scientific skills	5 marks
• Identifies problem and is capable of framing question. • Organizes and presents data in an organized fashion. • Student identifies discrepancies in data during discussions with laboratory group and some consultation with teacher. Students assume responsibilities for initiating discussions about scientific inaccuracies. • Identifies cause and effect relationships, resulting from data collection or quantitative analysis. • Independently, identifies sources of error, re-structures laboratory activity. Determines the reliability of the data. • Students identifies assumptions relating to measurement and/or analysis.	
• Scientific accuracy	5 marks
• Scientific terminology is used appropriately. • Student identifies and clearly states problems from independent reading, prior investigation, or case study. • Students look for trend or relationships in developing scientific theories and/or laws. Inductive thinking is stressed. • Text information serves as a springboard for further research. Proposed explanations and rationale for explanations is provided. Evaluates assumptions and the effect of bias. • Explains the problem investigated. Provides explanations and draws conclusions related to hypothesis. Modifies theory and/or hypothesis on the basis of results.	
• Communication	5 marks
• Portfolio is well organized with a table of contents and identified questions and answers. • Journal entries demonstrate metacognition. • The writers ideas are expressed in a purposeful and coherent manner. The focus is maintained throughout the presentation and transitions connect ideas appropriately. • Student initiated reconnaissance of experimental design and/or data collection. Student discusses the limitations of the experimental design or data collected. Evaluates assumptions and the effects of bias. Evaluates experimental design and re-structures experiment.	
• Conventions	5 marks
• Writing is essentially free from error in spelling, punctuation, and grammar. Errors that are present do not affect the clarity of presentation.	

Appendix #7: Lessons developed

7A: Investigation: Using a Microscope for Forensic Investigations

Police departments throughout the world use scientific instruments to help identify criminals. The microscope has played an important investigative role in the conviction or release of many individuals suspected of crime.

The Crime

Mrs. Argent, a wealthy stockbroker, was found dead in her swimming pool in the early hours of the evening. Red cotton fragments were found under her fingernails. The police believe that these fragments may have come from the murderer because Mrs. Argent was not wearing any cotton clothing. She was wearing a wool dress with a silk scarf. Three individuals were known to have see her that day. Cloth samples were taken from the clothing and from under the fingernails of the three suspects. Cloth samples taken from the victim's dress and scarf have been prepared to help you with your investigation.

The Suspects

Catch art:	Ms. Fleur, the gardener, had cloth samples taken from her blouse. She also had a pair of red pants with a tear in them. Ms. Fleur had been named in Mrs. Argent's will.
Catch art:	Mr. Maison, the butler, had a variety of cloth fibres found under his fingernails. His explanation was that he had moved carpets earlier that day. The fibres must have been from the polyester carpets.
Catch art:	Dr. Helper, the family physician, also had cloth samples taken from her blouse. The physician was a long-time friend of the victim, and had once dated her husband.

The Evidence

Cloth samples from Mrs. Argent's dress and scarf	Cloth sample taken from Ms Fleur's blouse and her torn pants
Cloth fibres found under Mr. Maison's fingernails	Cloth sample taken from Dr. Helper's blouse

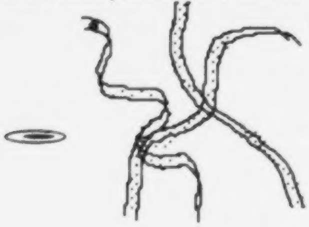
Materials

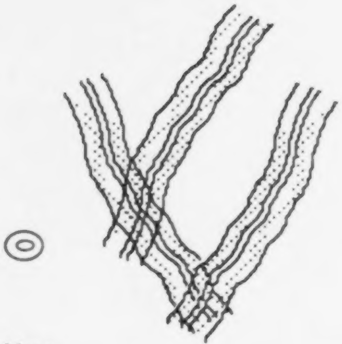
cloth samples taken as evidence
light microscope
microscope slides.
cover slips

Procedure

1. Design a plan to find the murderer.
2. Present relevant data that you have found in a data table. (You are required to make your own data table.) Consider the following factors:colour of material, type of material, size of fabric weave.

Cloth Samples under a Microscope



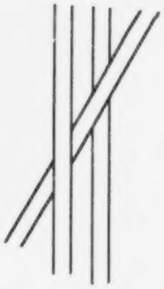


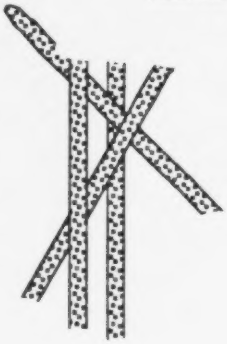
Cotton

- flat, ribbon-like structures
- often twisted or spiral with small bumps

Linen

- tube-shaped, even central canal
- thick outer walls with small bumps.



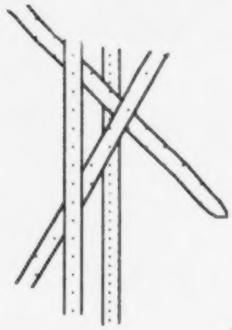


Polyester

- thick outer walls much like linen
- no central canal

Wool

- scaly, thick walls with no small bulges
- smooth, regular central canal
- regular, solid threads.
- like polyester, it has no central canal.
- granules often seen in the fibre.



Silk

Questions

1. Who do you believe murdered Mrs. Argent?
2. What evidence did you select in coming to your conclusion?
3. Does any evidence from this investigation provide conclusive prove? Explain your answer.

7B: Investigation: Writing a Mystery

Have you ever thought of writing a mystery novel? In this assignment you will be provided with additional evidence that you may select to write a mystery. Try the following things before beginning to write your murder mystery.

The assignment:

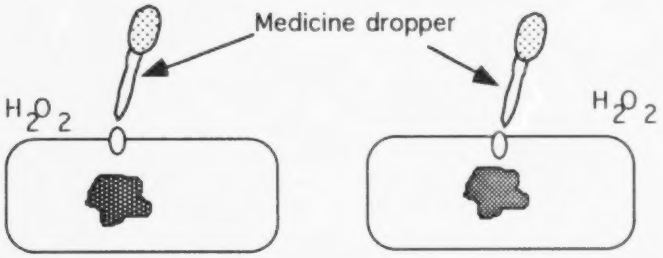
- You must set the scene for your murder mystery. Write a one page scene where you tell about the crime and profile the suspects.
- Provide evidence for investigators. Make sure that the evidence is clearly labeled.
- Once your laboratory group has completed writing the mystery, your teacher will give the mystery to another lab group to solve. In turn you will receive another mystery that your group will work to solve.

Materials:

Cloth samples containing blood and red dye.
Hydrogen peroxide
Compound light Microscope
Pencil
Ruler
Medicine dropper
50-ml beaker
Transparent adhesive tape
Small paint brush
Dissecting microscope or hand lens

Is the stain blood or just red dye?

- Blood stains can be identified by using hydrogen peroxide H_2O_2. Add a drop of hydrogen peroxide to cloth fibers. If it bubbles, living tissues are present. (We will use fish blood only for this investigation).

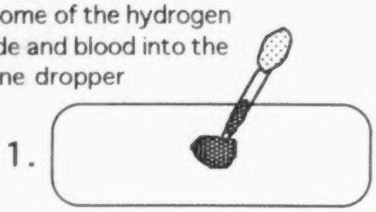


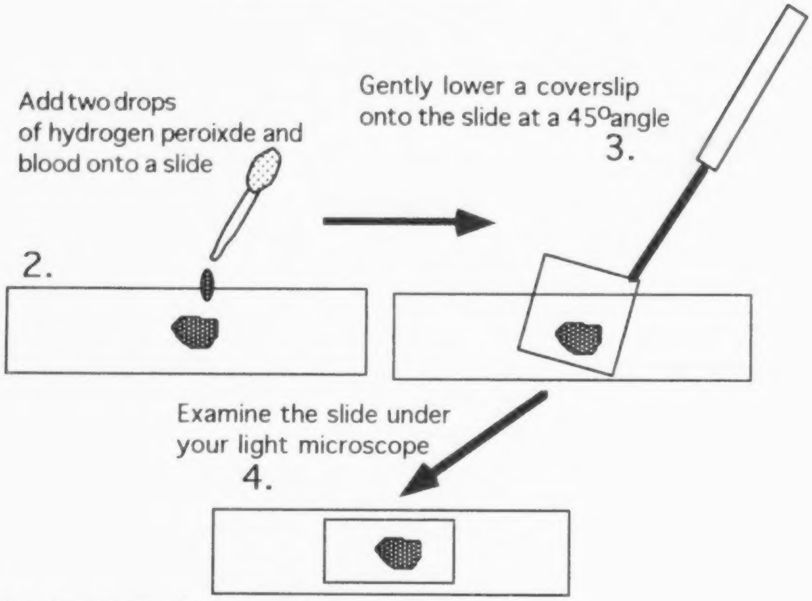
cloth with stain bubbles observed

cloth with stain no bubbles observed

- Place a drop of the hydrogen peroxide on a microscope slide and add a coverslip. Can you see any disk-shaped blood cells? Try medium and high power.

Draw some of the hydrogen peroxide and blood into the medicine dropper





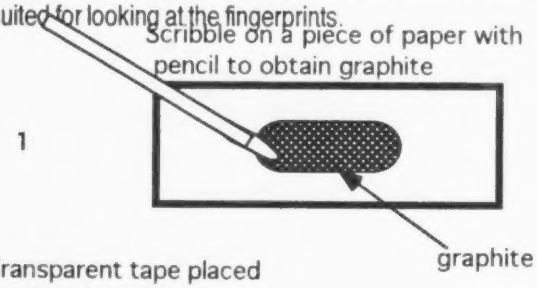
Add two drops of hydrogen peroixde and blood onto a slide

Gently lower a coverslip onto the slide at a 45°angle

Examine the slide under your light microscope

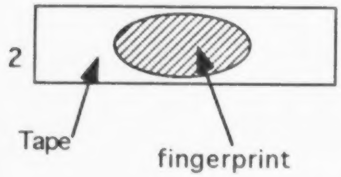
Are the fingerprints a match?

- Fingerprints can often be lifted by the sticky side of transparent tape. A light dusting with graphite often helps you see the fingerprints. A hand lens or dissecting microscope is suited for looking at the fingerprints.

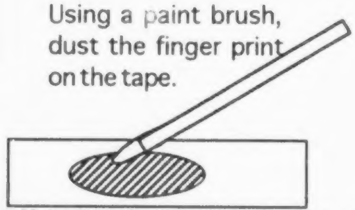


Scribble on a piece of paper with pencil to obtain graphite

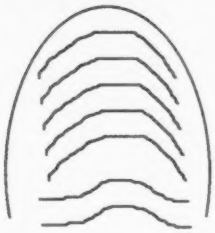
Transparent tape placed on top of the finger print



Using a paint brush, dust the finger print on the tape.

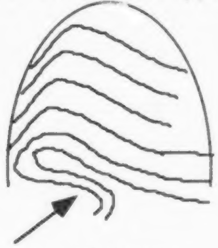


- Finger prints can be grouped into three large categories: the arch, the loop and the whorl.



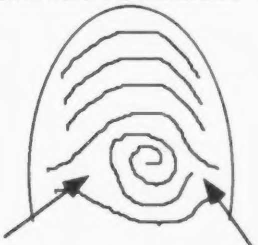
The arch

The arch is formed from ridges on the fingers that run from side to side with a slight curve. Note how all curves appear symmetrical.



The loop

The loop is formed from ridges on the finger that start from one side of the finger and curve back toward the same side. A small triangular area called the delta is formed.



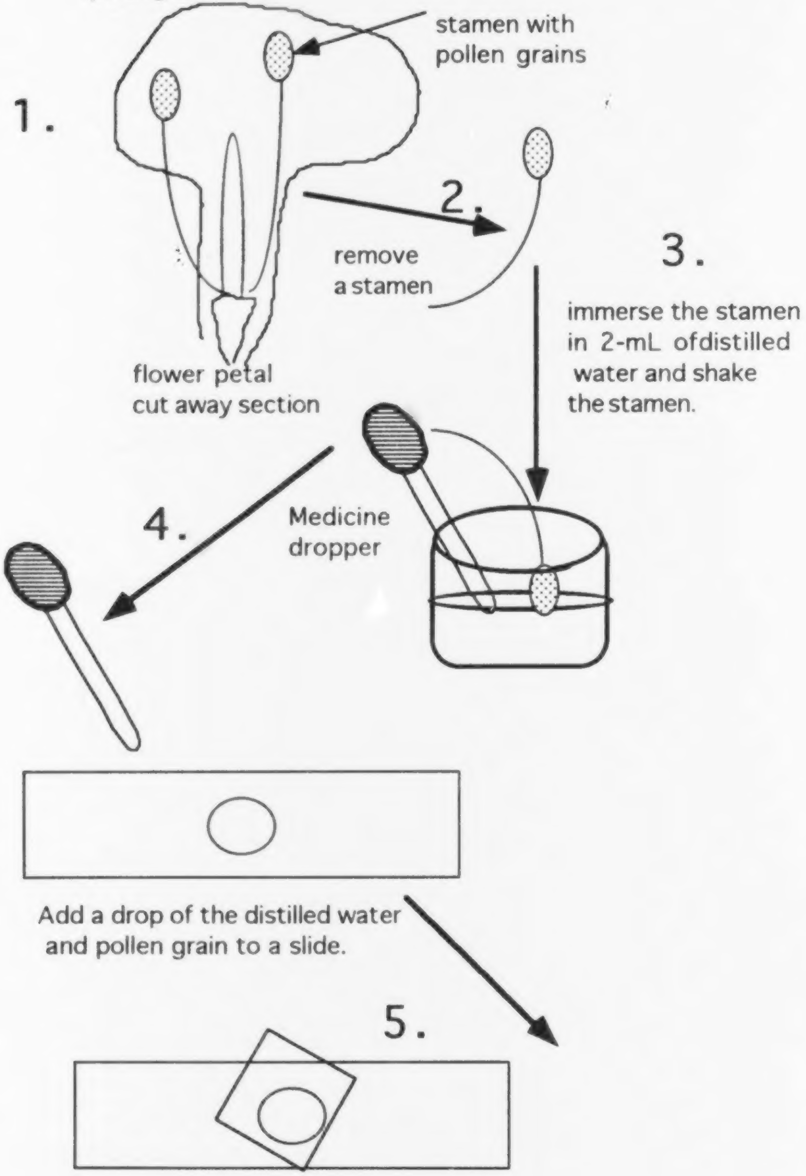
The whorl

The whorl is formed by a spiral that is found near the center of the finger. Two small delta's are found on either side of the whorl.

Determining location by testing for pollen

- Pollen, the male sex cells in plants, can provide important clues in solving a mystery. When viewed under a microscope, the distinctive pollen grains from different plants can be identified. Size, shape, color and texture are all important in establishing an identification.
- A murder suspect, who claimed to be far from the crime scene, was found to have pollen from a plant distinctive of that region in the wax of his ears. The man must have been in the area recently.

- Your teacher may have a number of prepared slides of pollen that you may use to build your investigation. If not, the following procedure will help you prepare slides of different pollen grains.


stamen with
pollen grains
1.
2.
remove
a stamen
3.
immerse the stamen
in 2-mL ofdistilled
water and shake
the stamen.
flower petal
cut away section
4.
Medicine
dropper
Add a drop of the distilled water
and pollen grain to a slide.
5.

Position a coverslip and view

- Just a recommendation: you might decide to provide a key with diagrams to identify various types of pollen that you are using for evidence.

Other things you may wish to consider:

- Matching carpet samples. The size of the carpet thread, the type of twist, and color are useful.

- Soil samples can be used. How much sand is in the soil. A soil test kit may be used to help this part of the investigation.
- Pictures of the etching on bullets. See the teacher's edition for various pictures.
- Hair color, and curl can be analyzed under the microscope.

Evaluation:

You will be given two evaluation forms. The first form is designed to analyze the work of your laboratory group as you worked through the murder mystery. The second evaluation form will allow you to assess the writers of the murder mystery that you attempted to solve. Is the mystery solvable?

- Outline some of the decisions made by your laboratory group as you attempted to write the mystery.
- Based on the feedback you received from the group that attempted to solve your mystery, what changes would you make?

Appendix #8: Journal Entries

- Students might be asked to critically reflect on Lovelock's, Gaia hypothesis. In what ways is a comparison of the earth to the human body limited.

- Scientific theories are often described as tentative explanations of natural phenomena. Theories are socially constructed explanations to open-ended questions. Ask students to find examples of open-ended questions in presented in the chapter. Why might scientists disagree about particular explanations.

- The city of Los Angeles is very concerned about emissions from automobiles. Automobile emissions have environmental implications for acid deposition and the greenhouse effect, as well as a number of related health problems. By the year 2000 automobile companies who wish to sell cars in Los Angeles must account for 10% of sales with non internal, combustion engines. Should your city follow Los Angeles' lead?

- Students may be asked to express concerns about ozone depletion, acid deposition or global warming. Do they believe that these problems really exist? Do they believe that any of these environmental problems affected their health? How would they go about changing things?

- Construct a concept map from chapter 1: Equilibrium in the Biosphere. The following terms may be useful: biotic, abiotic, biosphere, population, community, atmosphere, photosynthesis, biogeochemical cycle, and cellular respiration.

- Students may be asked to describe home recycling projects, or to devise a plan for recycling household refuse.
- Students may be asked to comment on their understanding of the laws of Thermodynamics. Why must energy be continually added to an ecosystem? Is energy destroyed within an ecosystem?

- Students may be asked to classify a number of ecosystems in their own schoolyard. How many micro ecosystems can they classify.

- Students may be asked to express concerns about artificial ecosystems. For example, do they have allergies? Do they feel that rugs in the school contribute to health problems? Do they believe that plants are really useful? How would they go about changing their artificial ecosystem?

- Construct a concept map from chapter 2: Energy and Ecosystems. The following terms may be useful: Autotrophs, heterotrophs, producers, consumers, food chains, pyramids of energy, pyramid of biomass, pyramid of numbers, and biological amplification.

- Students may be encouraged to outline their thoughts as they prepared for the debate. Where did they do their research? Did they change their mind? Did they find it difficult to investigate and debate science-related social problems?

- Students may be encouraged to read mythological explanations for the origin of life. How do mythological explanations differ from scientific explanations?

- Students may be asked to view figure 4.7 of Nelson Biology and comment on how embryological evidence supports the theory of evolution. Why are the early stages of development so similar? Do humans pass through an evolutionary ancestry as we develop?

- Ask students to comment on the theory that Pangea was once a supercontinent. Why might someone remain unconvinced?

- Students can be encouraged to express the difficulties they experienced in formulating an understanding of this challenging chapter. What things did they employ to aid them in constructing their knowledge. For example, some students may indicate that their understanding of one-directional blood flow was aided by figure 10.18. Other students may have attempted a series of their own drawings or constructed concept maps. The journal entry can provide students with more information on how they learn.

- Students might be asked to write a dialogue between the text and themselves, as they refute the idea that the transport of nutrients within a multicellular organisms requires a circulatory system. By acting as a *Devil's Advocate*, they can challenge their own learning and push understanding to a higher level.

- Group thinking and decision-making strategies may be recorded as students prepare for the debate. Did the students change their mind during the preparation for the debate? Students may even be asked record their initial positions about the social issue and to reflect upon these feelings after the debate has been completed. Did they change their mind after listening to opposing arguments. Should people who refuse to alter lifestyles that are dangerous to their health be permitted equal access to health care?

- Students might record reasons why material learned in the chapter was particularly relevant to them. For example, students interested in athletics might indicate how they will use information in the section *Adjustments of the Circulatory System to Exercise*.

- Students might record reasons why material learned in the chapter was particularly relevant to them. For example, students interested in athletics might indicate how they will use information on vital capacity, inspiratory reserve volume, and tidal volume.

- Students can be encouraged to express the difficulties they experienced in formulating an understanding of concepts presented in the chapter. What things did they employ to aid them in constructing their knowledge. For example, some students may indicate that their understanding of carbon dioxide transport by viewing figure 12.13. Other students may have attempted a series of their own drawings or constructed concept maps. The journal entry can provide students with more information on how they learn.

- Decision-making strategies may be recorded as each group prepares for the debate. Did the students change their mind during the preparation for the debate? Students may even be asked record their initial positions about the social issue and to reflect upon these feelings after the debate has been completed. Did they change their mind after listening to opposing arguments. Should governments ban the sale of tobacco products.

- Students might be asked to write a dialogue between the text and themselves, as they refute the idea that males are no longer necessary for human reproduction. By acting as a *Devil's Advocate*, they can challenge their own learning and push understanding to a higher level.

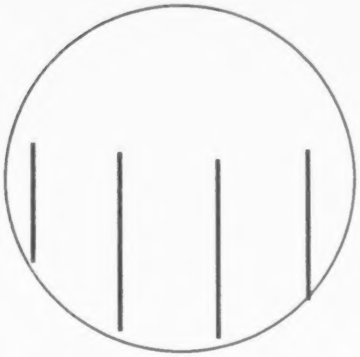
- Students can be encouraged to express the difficulties they experienced in formulating an understanding of this challenging chapter. What things did they employ to aid them in constructing their knowledge. For example, some students may indicate that they helped by the case study: *Hormone Levels During the Menstrual Cycle*. If completed by small groups, students might learn that active discussion is an effective way for them to learn. Other students may have attempted a series of their own drawings or constructed concept maps.

- Decision-making strategies may be recorded as each group prepares for the debate. Did the students change their mind during the preparation for the debate? Students may even be asked record their initial positions about the social issue and to reflect upon these feelings after the debate has been completed. Did they change their mind after listening to opposing arguments. Should laws be instituted to prevent pregnant women from drinking excessively?
- Students might be asked to write a dialogue between the text and themselves, as they refute the idea that virgin births are possible. They may consider what might happen if a primordial egg cell ovulates prior to meiosis. What would happen, should such an 2n-egg cell implant in the uterus. Some geneticists have indicated that the event, common in Daphnia and lower invertebrates that undergo parthenogenetic reproduction, might even occur in higher mammals. By acting as a *Devil's Advocate*, the students can challenge their own learning and push understanding to a higher level.
- Students can be encouraged to express the difficulties they experienced in formulating an understanding of this challenging chapter. What things did they employ to aid them in constructing their knowledge. For example some students may indicate that the laboratory: *Human Karyotypes,* aided their understanding of nondisjunction. Other students may have attempted a series of their own drawings.
- Decision-making strategies may be recorded as each group prepares for the debate. Did the students change their mind during the preparation for the debate? Students may even be asked record their initial positions about the social issue and to reflect upon these feelings after the debate has been completed. Did they change their mind after listening to opposing arguments. Should limits be placed on reproductive technology?
- Students may identify open-ended questions not answered in the chapter. For example, why do sperm cells produce four cells following meiosis, compared to only one viable egg cell? Although the text provides a description of the differences, no mechanism for the differences in cell division is offered. Why are cytoplasmic division different than female cytoplasmic divisions during meiosis. What advantages are served by restricting the number of egg cells produced by meiosis?
- Students can be encouraged to express the difficulties they experienced in formulating an understanding of heredity. What things did they employ to aid them in constructing their knowledge. For example, some students may indicate that the laboratory: *Genetics of Corn,* aided their understanding of dihybrid crosses. Other students may have attempted a series of their own summary charts or constructed concept maps.
- Decision-making strategies may be recorded as each group prepares for the debate. Did the students change their mind during the preparation for the debate? Students may even be asked record their initial positions about the social issue and to reflect upon these feelings after the debate has been completed. Did they change their mind after listening to opposing arguments.

Appendix 9: Performance-Based Assessment Strategies

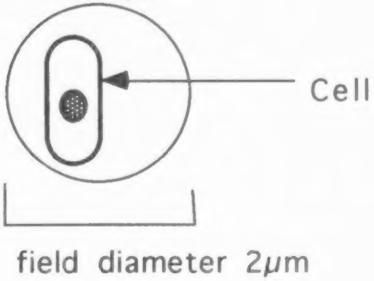
9A: Science 10: Microscope Work

1. A ruler was placed on the stage of a microscope and viewed under low power magnification. Each of the lines shown in the diagram below represent each mm marker. Indicate the field of view for low power magnification in mm and µm. (1 mm = 1000 µm).



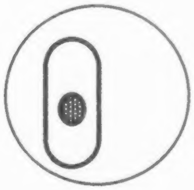
2. If the low power objective is 4X and the medium power objective lens is 10X, Calculate the field diameter for medium power. The field of view under low power magnification, for this particular microscope, was found to be 4 mm.

3. The diagram below shows a cell under medium power magnification. For the microscope used, it was determined that the field of view was 2000 µm. Calculate the size of the image.

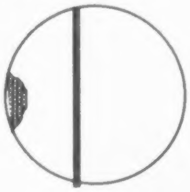


4. For the cell shown above, which diagram would best represent its appearance under high power magnification. Explain your reasons.

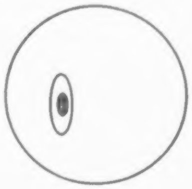
Possibility #1



Possibility #2



Possibility #3



9A: Assessment Template for the Microscope

check list (√)	Skill demonstrated	Comment

1.	Uses low power magnification to scan for objects.	
2.	Uses fine adjustment focus only when using high and medium power magnifications	
3.	Is able to prepare a dry mount.	
4.	Uses measurement units correctly (conversion of mm to μm).	
5.	Is able to determine field of view for low power magnification from observations	
6.	Is able to calculate field diameter for medium and high power magnifications from a ratio.	
7.	Can determine the size of an object using the microscope.	

9B: Science 10, Osmosis and the Concentration of Solute

Challenge

- How heavy can you make an egg? Work with a partner or in a group. Take the initial mass of an egg and immerse it in various solutions. Take the final mass of the egg. **Caution:** Check any procedures with your teacher before you begin.

Circle the yes if the outcome has been observed.

Prediction made	yes	no
Procedure is followed	yes	no
Balance is used correctly to measure the mass	yes	no
Data table made and results reported	yes	no
Graph made with mass plotted on X axis and time plotted along Y axis	yes	no
Title provided for graph	yes	no
Proper scale is used for X and Y axis	yes	no
Units provided for X and Y axis	yes	no
Line graph is used	yes	no
Use of a control	yes	no
Group demonstrates collaborative work environment	yes	no
Safety is observed during the laboratory	yes	no
Laboratory area is left clean and equipment is returned	yes	no
Conclusion provided	yes	no

12 to 13 yes responses, student mark = 5
9 to 11 yes responses, student mark = 4
6 to 8 yes responses, student mark = 3
4 to 5 yes responses, student mark = 2
less than 3 yes responses, student mark = 1
laboratory not done, student mark = 0

9B: Science 10, Performance-Based-Assessment for Challenge

Values: 1= not attempted, 2 = poor quality, 3 = average, 4 = very good

Category	Value				
1. Understands problems and uses technological thinking	1	2	3	4	comment
2. Is able to communicate plan and provide reasons for the design	1	2	3	4	comment
3. Mass of the egg	1	2	3	4	comment
4. Bonus marks for originality	1	2	Total =		

Rating scale for change of mass of the egg

4 Exceptional: The greatest change in mass of an egg recorded by a group. This provides the standard for comparisons.

3. Very Good: Change in mass within 20% of the standard.

2. Good: Change in mass between 40% and 20% of the standard.

1. Requires re-design: Change in mass between 80% and 40% of the optimum standard.

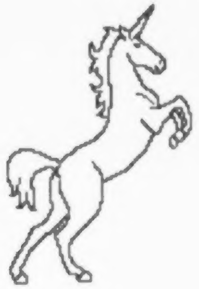
0. Faulty design: No change in mass or the egg lost mass.

9C: Science 14, Light Unit

Technology Link: Marketing and Colours

Your assignment:

You will use the picture below as part of a label for a company called Unicorn Light Bulbs. You want to convey a company image for strong soft light that is dependable. Only using colors and a label than contains no more than ten word, you must catch the consumers eye.



Instructions:

- Each group must work in isolation, no other group may see your design or know who has completed which design. Your teachers will give you group a code number as a means of identifying your design.
- Discuss ideas for color, font Type style), and the message provided on the product. No other images can be used.
- Your teachers will collect each of the designs and post them in the classroom.
- Record how each member of the group contributed to developing the label on a separate piece of paper. Once again, fix the code number identifying the group cooperation and turn it into your teacher.
- Each group will be provided with an judging form that will be used to determine which of the other groups has most eye-catching design. You will not rate your own group's design.
- Use the rating scale below to judge each of the designs for color and eye-catching words. #1= poor use of colors and words; #2 = Either weakness identified for the use of colors or the message; #3 = good use of colors and is eye-catching; #4 = exceptional (only one four may be issued).
- Turn in your scoring cards to your teacher. The scores will be tallied for each of the groups by your teachers and the top scores will be posted.

9C: Science 14, Performance-Based-Assessment
Judging group: (names of the group members doing the judging)_______________

Judging Card:

Code number	Score for the use of color	Score for the message	Total score

Teacher scoring

10 point = the group that finished first
8 points = groups that finished from 2 to 4
6 points = groups that finished from 5 to 7
(Notes this scoring guide maybe adjusted, depending on the size of the class and number of students in each group).
Notes to the teacher: Consider using the template for group assessment, found in the appendix of the ATE.

9D: Biology 30, Dynamic Model of the Cell

Although mitosis is described in stages, the process of cell division is continuous. To help you understand this process, work with a partner to build a dynamic model in which chromosomes can be moved to show the events of cell division. To keep your model simple, use only three chromosomes. In your model be sure that you are able to line the double-stranded chromosomes up in the centre of the cell, and that the single strands are able to move to opposite ends of the cell as they do during anaphase.

9D: Biology 30, Performance-Based-Assessment for The Dynamic Model of Cell Division

Group members:________

Part A: Detail and Accuracy of Design:__marks__

(maximum (10 marks)

Assign one mark for each of the features of the model to a maximum of 10 marks.

Descriptor	Yes	No
1. Chromosomes initially double-stranded becomes single stranded after the division.		
2. Students have devised a way of holding double-stranded chromosomes together (equivalent of centromere in actual cell).		
3. Model has a structure similar to the spindle fibers.		
4. Students have devised a method for attaching double-stranded chromosomes to the spindle fibers.		
5. Equatorial plate or equivalent is presented or can be identified in the model.		
6. Students have devised a way of separating the double-stranded chromosomes into single strands during anaphase.		
7. The single strands of the chromosome move to opposite poles during anaphase.		
8. Reorganization of two nuclei with single-stranded chromosomes.		
9. The dynamic model shows division of cytoplasm.		
10. Students devise a method for duplicating the single-stranded chromosomes.		
11. The model shows the formation of two new cell membranes following telophase.		
12. The models shows the formation of new nuclear membranes following telophase.		

Part B: Technological Thinking

Category	Value				
1. Understands problems and uses technological thinking	1	2	3	4	comment
2. Is able to communicate plan and provide reasons for the design	1	2	3	4	comment
3. Mechanics of the model (does it work)	1	2	3	4	comment
4. Bonus marks for originality	1	2	Total =		

9E: Biology 20, Designing an Experiment to Determine How Environmental Factors Affect Seed Germination.

A seed is something like a packaged plant. Everything the new plant needs is found inside of the seed. It contains the embryo and a packaged food supply. The protective coat provides resistance to cold and prevents drying for many months or in some cases even years. Each seed is specially adapted for specific environmental influences.

In this activity, you will design an experiment to determine how various environmental factors affect seed germination. Because not all seeds are the same, your conclusions must be restricted to those seeds you are studying.

Materials:

Per team of 3 to 4 students	For the class
Petri dishes (4 maximum)	Incubator or warm area
Paper towel cut to fit in Petri dish	Refrigerator or cooler with ice
10 radish seed (or substitute)	200 ml of 0.01 M acetic acid
10 tomato seeds (or substitute)	200 ml of 0.1 M acetic acid
10 lettuce seeds (or substitute)	200 ml of 1.0 M acetic acid
10 bean seeds (or substitute)	200 ml of 0.01 M NaOH
Safety goggles	200 ml of 0.1 M NaOH
Aprons	200 ml of 1.0 M NaOH
Medicine dropper	2, Thermometers
10 ml graduated cylinder	Light source
Forceps	Dark cupboard
Wax pencil or labels	Filter paper
Ruler (mm)	

Procedure:

1. Your teacher will organize the class into research teams. Each research team will be assigned a particular problem. You will be asked to design a laboratory procedure that helps you investigate the problem. You must submit a detailed laboratory procedure and have it approved for safety before beginning your investigation . Note some groups may be assigned one or two types of seeds only.

The research problems:

A. How does temperature affect the germination and growth of seeds? Consider testing in a warm environment, room temperature, and a cold temperature.
Group #1: Use Radish and tomato seeds
Group #2: Use Bean and lettuce seeds

B. Do seeds germinate and grow best when exposed to light? Consider using different light sources (artificial and natural) and a dark area.
Group #3: Use Radish and tomato seeds
Group #4: Use Bean and lettuce seeds

C. Do plants grow well in acidic conditions? Use the different acetic acid (vinegar) solutions. Hint: make sure you keep the volume of acid constant for each trial.
Group #5: Use Radish and tomato seeds
Group #6: Use Bean and lettuce seeds

D. Do plants grow well in basic conditions? Use the different NaOH solutions as the base. Hint: make sure that you keep the volume of base constant for each of the trials.
Group #7: Use Radish and tomato seeds
Group #8: Use Bean and lettuce seeds

2. Read the following hints before beginning to write your procedure.

Hint #1: Make sure that you wear safety goggles and a lab apron for the entire procedure. Even if your procedure does not require you to work with an acid or base, other people in the class will be. The entire workplace must be safe.

Hint #2: Make sure that your seeds do not dry out. Use lids for the Petri dishes.

Hint #3: Check your seeds for mold growth. Mold will slow growth and eventually kill the seeds. Use forceps to handle the germinating seeds.

Hint #4: Filter paper can be used as a divider.

Hint #5: Make sure you try to measure any changes that you observed. Experiments that can demonstrate the amount of difference are much more valuable for these types of questions.

3. Work with your group members to develop a procedure. Any group investigating a specific research problem, such as temperature, should check their procedure with other groups investigating the same problem. Laboratory data can be compared at the end of the laboratory.
   - Identify a control that you have used for the experiment.
   - Identify the manipulated (independent) and responding (dependent) variables.
   - List variables that you have controlled. (e.g.. the same amount of acid was used for each test.)
4. Present your procedure to your teacher and upon approval begin the experiment.
   - Construct a data table and record your results during following classes.

Questions

1. What conclusions could you draw from your experiment?
2. What other experiments might be needed to test your conclusions?
3. If possible present your data by way of a graph.

Application

4. Why might scientists be interested in answering the question that you researched?

9E: Biology 20, Assessment of Student Performance

Skills Assessment Total = 5 marks

1. Data tables are used (independent / dependent variables are identified)	yes	no
2. Units for measurement in data table are provides and appropriate	yes	no
3. Data collection is neat and organized	yes	no
4. Variables that must be controlled are identified in the procedure.	yes	no
5. Safety considerations are identified in the procedure.	yes	no
6. Line graph was selected as the appropriate graph to show changes in growth over time.	yes	no
7 Graph made with mass plotted on X and time plotted along axis	yes	no
8. Title provided for graph	yes	no
9. Proper scale is used for X and Y axis	yes	no
10. Units provided for X and Y axis	yes	no
11. Conclusion integrate the data collected	yes	no
12. Conclusions use qualifiers or limitations acknowledged.	yes	no
13. Appropriateness of design is evaluated	yes	no
14. Research is conducted to support findings. Research findings as linked with technological or environmental applications	yes	no

13 to 14 yes responses, student mark = 5

10 to 12 yes responses, student mark = 4

7 to 9 yes responses, student mark = 3

4 to 6 yes responses, student mark = 2

less than 3 yes responses, student mark = 1

laboratory not done, student mark = 0

Attitudes Total = 5 marks

Demonstrates collaborative work environment	yes	no
Initiates investigations and sustains involvement through the project	yes	no
Safety is observed during the laboratory	yes	no
Laboratory area is left clean and equipment is returned	yes	no

Provide a mark for each quality identified.

Technological thinking

Values: 1= not attempted, 2 = poor quality, 3 = average, 4 = very good

Total = 12 marks

Category	Value				
1. Understands problems and uses technological thinking	1	2	3	4	comment
2. Is able to communicate plan and provide reasons for the design	1	2	3	4	comment
3. Conclusions are insightful	1	2	3	4	comment
4. Bonus marks for originality	1	2	Total =		

Rating scale for report

4	Exceptional:
3.	Very Good:
2.	Good:
1.	Requires re-design:
0.	Faulty design:

9F: Chemistry 30: Antioxidants

A great deal of excitement is being generated by a group of vitamins--C, E, and beta carotene (the chemical parent of vitamin A). These chemicals are known as **antioxidants**. Early research suggests that these chemicals are able to make a group of harmful molecules, known as oxygen-free radicals, less dangerous.

The free radicals are created in your body by exposure to sunlight, X rays, ozone, tobacco smoke, car exhaust, and other environmental pollutants. They damage the genetic information in your body cells, causing mutations. The altered instructions found in the genetic information can cause the cell to divide at uncontrolled rates or even die. Scientists believe that these free radicals play a major role in the development of cancer, heart or lung disease, and even cataracts (a condition that makes the lens of the eye to become opaque). By taking these chemicals "out of commission", you would live longer and experience better health.

- Imagine getting the contract to design an advertising campaign for antioxidants. Design an advertising poster for these vitamins. Be prepared to support any health claims made on the poster. A summary list of support for your claims should be submitted along with the poster.

9F: Chemistry 20, Assessment Template Antioxidants

Students name:__

Group members:________

Levels of Attainment

Descriptors		poor	average	good	excellent
1.	Evidence of research about antioxidants • bibliography compiled or, • use of periodical index or, • use of electronic research	1	2	3	4
2.	Knowledge is expressed in poster • vitamins C, D, E and beta carotene identified • foods containing these vitamins identified • oxygen-free radicals described • anti-aging claims • links of free radicals to cancers	1	2	3	4
3.	Design and appeal • quality of diagrams • combination of eye-catching colors • organization of information conveys the intended message	1	2	3	4
4.	Field test of design proposals • survey of poster was developed • alternatives were considered	1	2	3	4

Total score: maximum 16 points

9G: Chemistry 20, Determining Vitamin C Levels in Juice Drinks

Have you ever read a consumer report? Many products claim to be vitamin enriched. How much vitamin C is present in orange drinks? In this investigation you will have an opportunity to perform consumer tests for vitamin C on different types of fruit drinks.

Materials

safety goggles
lab apron
6 test tubes
wax pencil
10-mL graduated cylinder
indophenol
medicine dropper
orange juice
distilled water
selected fruit drinks such as orange drink, C-Plus orange drink, orange crystals, white grape juice, lemon juice

Procedure

1. Put on your safety goggles and lab apron.
2. Label one test tube "C" for control and another "T" for test.
3. Using a graduated cylinder add 5 mL of indophenol to both test tubes "C" and "T".
    - Record the initial colour of the indophenol.
4. Using the medicine dropper add 7 drops of orange juice to test tube "T". Shake the test tube and observe for a colour change. If no colour change occurs, continue adding single drops of orange juice, shaking the test tube immediately after adding each drop.
    - Record the number of drops required to have the indophenol indicator become colourless.

Catch art: Figure 16.8.2

5. Add the same number of drops of distilled water to the indophenol indicator in the test tube labelled "C".
    - Record your observations.
6. Devise a method for converting the number of drops added to the test tube containing indophenol solution to a measurement in millimetres. (Refer to Skills on measuring liquids on page 00.)
    - State the method you used to measure the number of drops in millilitres.
7. Work in groups. Repeat the procedure described above to test the other solutions for vitamin C.
    - Prepare your own data table showing the number of drops of juice required to turn the indophenol indicator colourless. Convert the number of drops to millilitres.

Questions

1. Which of the drinks contained vitamin C?
2. Which drink contained the greatest amount of vitamin C?
3. Prepare a bar graph showing the levels of vitamin C found in various drinks.

Apply

4. Would you recommend orange juice over orange drink? Give your reasons.
5. If the person's diet already contained a rich supply of vitamin C, could any harm occur if vitamin C supplements were taken daily? Explain your answer.

Extension

6. Devise an investigation to compare the vitamin C content of freshly squeezed orange juice to orange juice that has been stored in the refrigerator for a long period of time. Check your procedure with your teacher, and carry out your investigation.
7. Test other foods for vitamin C.

9G: Chemistry 20, Assessment Template Determining Vitamin C Levels in Juices and Drinks

Students name:__

Group members:________

Use the following checklist as a guide for assessment.

Descriptor	yes	No
Laboratory Safety • goggles and apron used correctly for the entire laboratory • Indophenol solution does not make contact with skin surfaces • all glassware is washed and the work area is left clean • proper disposal of chemicals following the laboratory.		
Laboratory Technique • test tube with different drinks are clearly labeled • measurements of drinks are done accurately with a graduated cylinder or pipette • graduated cylinders and medicine droppers are rinsed with distilled water after each transfer of different drinks • thumb not placed over the test tube when shaking solutions • each member of the group is active and involved • a method of converting the number of drops to a volume measurement is completed with accuracy		
Analysis and Communication • members of the group are able to construct their own data tables. • dependent and independent variables are identified in the data table. • the number of drops of solution added is recorded • the number of drops added is converted to a volume measurement and presented in data table • all relevant data in presented in an easy to read table format • accurate reporting of the laboratory data • Plausible conclusion is provided		
Graphing Skills • a bar graph is used • units of measurement found on X and Y axis. • graph has a title • independent variable is presented on the X-axis • dependent variable is presented on the Y-axis • units are provided with a scale on both the X and Y axes.		

Conversion scale for scoring

20 or more (√) marks = 10 points
18 to 19 (√) = 9 points
16 to 17 (√) = 8 points
14 to 15 (√) = 7 points
12 to 13 (√) = 6 points
10 to 11 (√) = 5 points
8 to 9 (√) = 4 points
6 to 7 (√) = 3 points
4 to 5 (√) = 2 points
1 to 3 (√) = 1 point
Not complete = 0

10: Table of inventions

Note: The horizontal columns identify personal interactions while the vertical columns identify the group interactions.

	Teacher (a) colleagues	Students (b) collective	Subject matter (c) (Science)	Milieu (d) Community
Teacher (1) Individual	1. What are teachers' interpretations of the program of studies? 2. What are the multiple definitions of STS? 3. Will teacher autonomy and professionalism be compromised by a collaborative project? 4. Are leadership roles already established?	1. Will the STS program initiate broadened forms of student evaluation and change testing? 2. Will changes in evaluation strategies change the relationship between teacher and student. (i.e portfolios may reduce teacher role as judge of student work, and students may become more actively involved in their own assessment)?	1. Will teacher's be able to provide a different view of science. (i.e. it provides objective, factual answers. Subjective evaluation may also show the subjective nature of science and the tentativeness of scientific theories. 2. Can aspects of STS science be evaluated? 3. Will teachers accept and value new assessment strategies?	1. Will the administration, and community accept a movement from objective to more subjective evaluation methodologies? 2. Will definitions and understandings of what constitutes subjective and objective evaluation change?
Students (2) individual	1. How will classroom dynamics change? 2. Will classrooms become more democratic? 3. Will alternative evaluation strategies change the perceptions about what is valued?	1. Will the competition emphasis be altered? 2. Will student leaders emerge?	1. Will subject material become personally relevant? 2. Will STS science be valued? 3. Will students accept and value new assessment strategies?	Will students respond positively to new strategies and public perceptions of how these strategies impact student understanding?
Subject (3) subjects other than science	1. Will movement toward the types of assessment strategies used in humanities help bridge the gap? 2. Will humanities teachers provide a leadership role in helping science teachers to explore unfamiliar assessment strategies?	1. Will students see greater connection between science and non-science subjects?	Will collaborative opportunities be presented among disciplines?	Will the community see greater connection between the evaluation practices across the curriculum?
Milieu (4) school	Will science teachers, administrators, and teachers from other disciplines communicate about evaluation.	Will school atmosphere change?	Will cooperation, collaboration, and peer support be more accessible?	How will be the school be perceived by the community?